The Journal of Experimental Education

A periodical report of scientific investigations relating to child development, curriculum, learning, teaching, supervision, measurements, statistics, and experimental techniques.

Volume XV | DECEMBER, 1946 | Number 2

CONTENTS

PAGE

Relationships Between Teacher Certification and Education in Wisconsin: A Study of Their Effects on Beginning Teachers: *Lois Gadd Nemec* 101

Factors Affecting Regularity of the Flow of Words During Written Composition: *John Andrew Van Bruggen* 133

A Study of Children's Heroes and Ideals: *M. Louise Stoughton and Alice M. Ray* 156

Children's Attitude Toward Discipline as Related to Socioeconomic Status: *Laura Dolger and Janet Ginandes* 161

The Effect of Training on the Tempo of Development of Graphic Representation in Preschool Children: *Elisabeth Ruch Dubin* 166

$5.00 A YEAR | PUBLISHED QUARTERLY | $1.50 A COPY

Published by Dember Publications, Inc., 114 S. Carroll Street, Madison 3, Wisconsin

Entered as second-class matter October 17, 1938 at the post office at Madison, Wisconsin, under the act of March 3, 1879.

EDITORIAL BOARD

A. S. Barr, Chairman, Professor of Education, University of Wisconsin, Madison 6, Wis.

Arthur T. Jersild, Professor of Education, Teachers College, Columbia University, New York City. Editorially responsible for materials on child welfare, guidance, and development, published each December.

Palmer O. Johnson, Professor of Education, University of Minnesota, Minneapolis, Minnesota. Editorially responsible for materials on measurements, statistics, and methods of experimental research, published each March.

H. H. Remmers, Professor of Educational Psychology, Director Division Educational Reference, Purdue University, Lafayette, Indiana. Editorially responsible for materials on learning, teaching and supervision, published each September.

J. Wayne Wrightstone, Assistant Director, Bureau of Reference, Research and Statistics, Board of Education of the City of New York, 110 Livingston Street, Brooklyn, New York. Editorially responsible for materials on curriculum construction, published each June.

CONTRIBUTING EDITORS

Emmett A. Betts, Research Professor and Director of Reading Clinic, Pennsylvania State College, State College, Pennsylvania.

Gilbert L. Betts, Supervisor of Graduate Research in Education, Colorado State College, Fort Collins, Colorado.

Leo J. Brueckner, Professor of Education, University of Minnesota, Minneapolis, Minnesota.

Oscar K. Buros, Associate Professor of Education, Rutgers University, New Brunswick, New Jersey.

Guy T. Buswell, Professor of Educational Psychology, University of Chicago, Chicago, Illinois.

Otis W. Caldwell, General Secretary, The American Association for the Advancement of Science, Boyce Thompson Institute for Plant Research, Yonkers, New York.

Leslie L. Chisholm, Associate Professor of Education, State College of Washington, Pullman, Washington.

Herbert S. Conrad, Technical Consultant, College Entrance Examination Board, Princeton, New Jersey.

Stephen M. Corey, Professor of Educational Psychology, University of Chicago, Chicago, Illinois.

Edward E. Cureton, Secretary–Treasurer, Richardson, Bellows, Henry & Company, Inc., New York 4, New York.

Robert A. Davis, Professor of Education, Director of Bureau of Educational Research, University of Colorado, Boulder, Colorado.

Harl R. Douglass, Director of College of Education, University of Colorado, Boulder, Colorado.

Harold A. Edgerton, Director, Occupational Opportunities Service, Professor of Psychology, Ohio State University, Columbus 10, Ohio.

John C. Flanagan, Professor of Psychology, University of Pittsburgh, Pittsburgh, Pennsylvania.

Carter V. Good, Acting Dean, Teachers College, University of Cincinnati, Cincinnati 21, Ohio.

Robert W. B. Jackson, Assistant Professor of Educational Research; Assistant Director, Department of Educational Research, Ontario College of Education, University of Toronto, Toronto, Canada.

Harold E. Jones, Professor of Psychology and Director, Institute of Child Welfare, University of California, Berkeley 4, California.

Noel Keys, Professor of Education and Lecturer in Human Relations, University of California, Berkeley, California.

D. Welty Lefever, Professor of Education, University of Southern California, Los Angeles, California.

Edward A. Lincoln, Consulting Psychologist, Halifax, Massachusetts.

Irving Lorge, Professor of Education, Executive Officer, Institute of Psychological Research, Teachers College, Columbia University, New York 27, New York.

A. R. Mead, Director of Educational Research, University of Florida, 330 P. K. Yonge Building, Gainesville, Florida.

T. E. Newland, Lt. Comdr., USNR, 2702 Wisconsin Avenue, N. W., Washington 7, D. C.

C. W. Odell, Professor of Education, University of Illinois, Urbana, Illinois.

Willard C. Olson, Professor of Education, Director of Research in Child Development, University of Michigan, Ann Arbor, Michigan.

W. E. Peik, Dean and Professor of Education, University of Minnesota, Minneapolis, Minnesota.

S. L. Pressey, Professor of Educational Psychology, Ohio State University, Columbus, Ohio.

Clarence E. Ragsdale, Professor of Education, University of Wisconsin, Madison, Wisconsin.

William Reitz, Associate Professor of Education, College of Education Examiner, Wayne University, Detroit 2, Michigan.

Henry D. Rinsland, Professor of Education and Director of Educational Research, The University of Oklahoma, Norman, Oklahoma.

Robert T. Rock, Jr., Professor of Psychology, Head of Dept. of Psychology, Graduate School, Fordham University, New York City.

Phillip J. Rulon, Professor of Education, Harvard Graduate School of Education, Cambridge 38, Mass.

Douglas E. Scates, Associate Professor of Education, Duke University, Durham, North Carolina.

Harold Seashore, Director, Test Division, The Psychological Corporation, New York 18, New York.

David Segel, Educational Consultant, Specialist in Tests and Measurements, Federal Security Agency, U. S. Office of Education, Washington, D. C.

Paul W. Terry, Professor of Educational Psychology, University of Alabama, University, Alabama.

Helen Thompson, Associate Attending Psychologist, New York Post–Graduate Hospital, 303 East 20th Street, New York 3, New York.

Robert L. Thorndike, Associate Professor of Education, Teachers College, Columbia University, New York City.

Herbert A. Toops, Professor of Psychology, Ohio State University, Columbus, Ohio.

Maurice E. Troyer, Director, Bureau of School Services, Syracuse University, Syracuse 10, New York.

Helen M. Walker, Professor of Education, Teachers College, Columbia University, New York City.

Beth L. Wellman, Professor of Psychology, Child Welfare Research Station, State University of Iowa, Iowa City, Iowa.

Guy M. Wilson, Professor of Education, Emeritus, Boston University, Boston, Massachusetts.

Paul A. Witty, Professor of Education, Director of Psycho-Educational Clinic, School of Education, Northwestern University, Evanston, Illinois.

Ernest R. Wood, Professor of Education, New York University, New York City.

D. A. Worcester, Chairman Department of Educational Psychology and Measurement, University of Nebraska, Lincoln, Nebraska.

DEMOCRAT PRINTING COMPANY
MADISON, WISCONSIN

Journal of Experimental Education

Volume XV | December, 1946 | Number 2

RELATIONSHIPS BETWEEN TEACHER CERTIFICATION AND EDUCATION IN WISCONSIN: A STUDY OF THEIR EFFECTS ON BEGINNING TEACHERS

Lois Gadd Nemec
University of Wisconsin

CHAPTER I

INTRODUCTION AND BACKGROUND

Introduction

Questions regarding teacher certification in Wisconsin, have grown more numerous and perplexing each year. Present practices are both commended and condemned by teachers, supervisors, administrators and teacher-training institutions. In any case there is some misunderstanding and some apprehension regarding the recommendations which are necessary for licensing beginning teachers. While some teachers are worried about the outcome of their probationary period of teaching, the state records show that many do not know there is such a thing as "being on probation" or even that they have to be certified, until they get into difficulty about it.

Certification and teacher-education cannot be studied by isolating one from the other inasmuch as the teacher-education institutions which give pre-service education must recommend their graduates for their first license to teach; and the principals, supervisors and superintendents, who continue—or should continue—with an in-service education program must recommend the beginning teachers for succeeding licenses and certificates.

The significance of the problems involved in this relationship between certification and teacher-education, and their effects on teaching personnel — both professionally and as human beings—have prompted this investigation.

Statement of the Problem

Accordingly the main purpose of this study is to derive recommendations for the improvement of teacher certification and education in Wisconsin from an examination, analysis and study of the state records of teachers who have had difficulty with certification.

There is a real challenge for professional people in a study of this sort, since difficulty with certification, and all its attendant problems, can gravely affect a teachers' morale and therefore his entire social and professional life. Whatever affects a teacher's work and personality is bound to have consequences for the child in his care; accordingly anything which can be done to improve the legal and professional services to teachers has a more far-reaching effect upon the educational program.

Teachers who have difficulty with certification, as defined in this study, fail to receive a state certificate at the end of the usual probationary period of two years, because one or more supervisory officials report unfavorably to the state superintendent regarding them. It is felt that an intensive study of the records of such teachers can assist in pointing out some of the relationships of and some of the deficiencies in the teacher-education and certification programs, and become the basis for recommendations for improvement.

Answers to the following questions will be sought:

1. What are the sources, causes, and results of the failure of some teachers to have certificates granted to them at the end of the usual probationary period?
   a. What are some of the problems these beginning teachers have, according to the records on file in the State Department of Public Instruction?
   b. Can any of the causes of these difficulties be traced to the pre-service education programs in our normals, colleges, or universities?

c. Are the criteria used to evaluate the work of these teachers, or the interpretation of the criteria by the people "rating" them, any source or cause of their failure? That is, is the in-service education program in any way responsible?

d. Could any of the sources of the difficulties lie in the communities to which these candidates go for their first years of teaching?

e. Are all of these teachers really failures?

f. What are the apparent results of the temporary or permanent failure of these teachers to qualify for a certificate? Are there any intangible results which may be as important, even though they cannot be handled statistically?

2. What are the effects of present methods of certification in Wisconsin, in its legal and administrative aspects, upon teacher education and the teachers themselves?

3. What are the implications of these questions and their answers for the improvement of certification and teacher education in Wisconsin?

Assumptions and Hypothesis

Some assumptions on which this study is based, are:

(1) That some sort of legal certification of teachers is necessary, and that a probationary period is desirable.

(2) That teacher education is a continuous process which is only begun in college and which needs to be extended through many if not all the years of teaching.

(3) That the records from the files of the State Department of Public Instruction which are used as a basis of the study are reliable and bear directly on the problem in hand. It is *not* maintained that all the evidence is objective, for the ratings and recommendations which made these teachers "cases" are based on subjective evaluations. This subjective evidence, however, will be dealt with as objectively as possible. After all, what happens to these teachers is just as real as though the evidence *were* objective.

The hypothesis here advanced is that a study of the records of teachers who fail to receive their certificates at the end of the two year probationary period, will give some bases from which to derive conclusions and recommendations for the improvement of the services of college faculties, supervisory officials, and the certification department, to beginning teachers.

Background and Limitations of the Study

Before 1936, teachers in Wisconsin could obtain an unlimited or "life" certificate after graduation from a two-year teachers college course and two years of successful teaching while holding two one-year licenses. Teachers graduating from a one-year course offered by a teachers college or a county normal, could be certified by a county or city superintendent, although these certificates were good only in the local community where issued.

In 1936, the state superintendent, who had the legal authority to do so, ruled that in the future no life or unlimited certificates would be granted until a four year course was completed. It was in this year that the five-year non-renewable certificate took effect. This latter certificate is issued to people who graduate from a two or three-year course and who have taught two years successfully on two one-year licenses.

Since 1939, when the law was passed which put *all* certification of teachers in the hands of the state superintendent, every teacher-candidate must receive a one-year license from the state department upon recommendation of a teacher-training institution. If this candidate is successful during the first year of teaching (that is, if he is recommended by his principal, supervisor, or superintendent) he is issued a second one-year license. If, at the end of his second year of teaching, he is again recommended by his school officials, he is granted a certificate by the state superintendent. This certificate is (1) a non-renewable certificate, good for only 5 years, if the candidate has completed a two or three year course at college; or (2) a life, or unlimited certificate, if he has completed a four year college course and thus earned a bachelor's degree. The two years of teaching before the issuance of a certificate constitute the "probationary period," and most teachers qualify for a certificate at the end of it.

The *Thirty First Biennial Report of the Superintendent of Public Instruction of the State of Wisconsin* (July 1942 to June 30, 1944) has this to say regarding "Teacher Training and Certification" (pages 25–26):

"Under Section 14.57 (22) the State Superintendent of Public Instruction is empowered to license and certificate teachers. Section 14.57 (22) reads as follows:

"'He shall license all teachers of the public schools of the state, and make rules and regulations and prescribe standards of attainment for the examination, licensing and certification of teachers.'

"There is no other certificating authority. Since July 1, 1939 licenses and certificates have been based upon diplomas granted by approved institutions at graduation. No general or blanket teaching licenses or certificates are issued. All licenses and certificates are based upon and determined by these specific types of preparation:—rural elementary schools; village and city elementary schools; secondary schools, and for the special fields.

"Rural teachers are recruited from the rural departments of the nine state teachers colleges and twenty-two county normal schools. Courses are two years in length and in the main are uniform. Graduates receive licenses and non-renewable certificates limited to a total of seven years and to this type of school. Additional preparation is required before an extension of time is granted.

"Practically all of the elementary teachers for village and city schools are graduates of three and four year courses in the state teachers colleges. For the three year course the license and certificate are limited by practically the same rules which apply to graduates from the two year rural course. Four year graduates receive a bachelor's degree from the teachers colleges and are granted life certificates after two years of successful teaching experience. Said certificates are valid only in elementary schools.

"To teach in high school, graduation from an accredited four years college or university course is required, with one academic major and two minors totalling 54 semester hours or two majors of twenty-four hours each. In addition they must have had eighteen hours of professional subjects, including practice teaching. Qualified graduates receive licenses limited to their fields of preparation. The University of Wisconsin, nine state teachers colleges, and thirteen private colleges in Wisconsin constitute the accredited list for high school teachers of academic subjects. Stout Institute is recognized for training teachers of home economics and industrial arts."

This brief summary of the modern history of teacher certification in Wisconsin points out some of the significant relationships between certification and teacher education, and explains some of the limitations of the study.

It explains, for one thing, the limitations in *time*. The year 1936 has been chosen for the first grouping because that is the year that the requirements for a life certificate were first raised to a four year college course, and the five-year non-renewable certificate was first issued. Any of the cases under consideration in this study who had licenses or certificates issued before that date are included in the "1936 and before" column in Table I. Another reason for choosing this time scope is that it includes the period of time in which records of the state certification office are kept in the "active files".

Another limitation should be pointed out here, although it will be dealt with again in the next chapter. The records examined showed that many teachers had difficulties, not only in their probationary period, but later, and sometimes after a life certificate had been granted. In fact, some failed for professional or personal reasons after apparently having no difficulties, or only minor ones, during the regular probationary period. These failures had nothing to do with certification requirements, however. Some other teachers, who did have trouble with certification, failed to acquire either a five year non-renewable or an unlimited certificate, because of some legal or technical difficulty, and not for any professional or personal reasons. When the "ruling" on life certificates was changed in 1936, and the law was changed in 1939, many teachers had trouble with license and certification requirements during the transition period. Cases of these kinds are not considered in this study.

Also excluded from consideration are teachers who, the records show, were marked below average and still received their certificates; reasons for this—such as a strong protest—were included in some records; in other cases no reasons are given. To be a "case" for this study, then, a teacher's record had to show first, that he was marked below average

TABLE I

NUMBERS OF CANDIDATES WHOSE LICENSES WERE ISSUED OR CONTINUED AND WHOSE CERTIFICATES WERE ISSUED OR DENIED IN THE TEN YEAR TIME SCOPE OF THE STUDY

Date of:	1936 and before	1937	1938	1939	1940	1941	1942	1943	1944	1945	Total
Issue of First License	72	26	23	42	23	38	28	13	0	0	265
Issuance of Last License	1	16	25	15	24	44	30	47	48	15	265
Issuance of Life Certificate	1	0	13	23	12	22	41	16	30	20	178
Denied a Life Certificate	0	0	3	1	1	1	3	13	15	1	38
License Continuation	0	0	0	0	0	1	0	3	34	11	49

on his reference blanks after he had applied either for another license or a certificate; and second, that he had been denied this license or certificate by the state department.

Lastly, the first part of this paper is limited to a description, tabulation and analysis of the data. No justification is offered for any particular practice or requirements. If the evidence seems to warrant it, recommendations for maintaining, correcting, or changing any procedures or requirements will be made in due course.

DESCRIPTION OF PROCEDURES AND SELECTED CASES

First of all, the cooperation of the State Superintendent and the Certification clerk in the Department of Public Instruction was sought and obtained. They generously granted full and free use of the records in the active files of the certification division. These active files contained the records of teachers who have applied for their first or second license, or one of the two types of certicates, in approximately the past ten years.

In all, 6,136 records, contained in 22 files, were carefully examined. As each record was lifted from the file, the reference or recommendation blanks were scanned first.* If marks were noted in the "below average" column, the record was next checked for the carbon copy of a letter from the state superintendent, saying that the candidate's request for a certificate was being denied. If such a letter was found, the record was removed from the file so that all data could be copied and tabulated.

Every such letter which accompanied these records, was worded almost exactly like the one quoted below, and was stamped with the signature of the state superintendent:

> "I am enclosing another one-year license, which will permit you to teach in Wisconsin schools for another year. The records on file in this office do not warrant the issuance of a certificate to you at this time. You are marked below average in 'Ability to Instruct', 'Ability to Control' and 'Cooperation'. Unless you improve considerably during the next year, I cannot issue you another license or a certificate. I trust you will make better progress during the coming year."

Two hundred sixty-five cases were found, each meeting the above requirements. Note that this is not a sample of the deviants—it is *all* of them. In other words, every record in the active files was examined and *all* cases having this type of difficulty to date with certification were selected for study. Therefore, no problem of sampling is involved. While this number is small, and only approximately 4½% of the total of 6,136 records in the active files, several points should be considered.

(1) In the first place these are supposedly selected people. The colleges and normals which graduated them recommended them as being ready and able to teach.

(2) Only one of the 265 cases was denied any further license after the first year of teaching. 264 teachers did not become "cases" until after their second year of experience. The 6,136 records include those of all teachers in the state who were just finishing their first and second years of teaching in the spring of 1945 when the tabulation was being completed. Therefore, the 4½ percentage is not really a limited figure, for it might

* A sample reference blank is included in the appendix, of original thesis on file, Library, University of Wisconsin.

easily be that some of the rest of the 5,871 teachers could have become "cases" a few weeks or months after this study was completed.

(3) The type of difficulty here being studied is not the only type of difficulty, trouble or failure that teachers have. Therefore, we can not, by any manner of means, say that only 4½% of some six thousand teachers failed, temporarily or permanently, and that this is too small a percentage to have any significance.

After these 265 cases were taken from the files, all possible information regarding them was tabulated on large sheets. Each case was given a number and his record was made as complete as possible. Valuable information about his background was obtained from the application blank on which the candidate applied for his certificate.*

From the large data sheets, the information was transferred to work tables, so that all possible comparisons and differences could be noted. Frequently some columns or rows of the work tables were combined in order to treat them statistically. Appropriate statistical treatment was then applied in order to test for significant differences, correlations, and other values. Finally, interpretations, conclusions, and recommendations were derived from the examination and analysis of the data and their treatment.

Detailed Description of the Selected Cases

It has already been pointed out that a candidate was not accepted as a "case" unless the record showed first, that he was marked below average on the recommendation blanks; second, that a letter had been written to him from the state department, telling him either that he was being denied another one-year license, or that he was being denied a certificate. Actually, only one of the 265 cases was denied a second one-year license; all others were given a second license and at least another year in which to "make good", even though their ratings were below average after the first year of teaching.

This might be called, in one sense, a study of failures, for all of these people except one failed in obtaining their certificates at the end of the ordinary probation period of two years.[1] However, the majority of them (178 of the 265) were finally granted a certificate after an extended probationary period of from one to five years; that is they acquired a certificate after teaching from three to seven years instead of the usual two.

Therefore, all of these people are not ultimately failures as far as certification is concerned. Instead, some of them are ultimately successes, not only in acquiring a certificate but in actual professional achievement. For it is known that a few of the cases that were at first denied, and then granted a certificate, have later become outstanding successes in the field of education.[2]

We cannot then, actually call this a study of failing teachers, because so far only 38 of the 265 have actually failed and been denied any legal right to teach in Wisconsin. Therefore, this is really a study of those teachers who had or are having *difficulty with certification and who were denied a certificate at least once by the state superintendent,* because they were marked below average by supervisory and administrative officials. This is the one factor that has been held constant in this study, and is the only factor common to each and every case.

The data for the figures just mentioned are tabulated in Table II, which shows the number of licenses granted before certification was either (1) achieved or (2) definitely denied, and the number of licenses issued up to the present time to those candidates whose certificates are (3) still pending. Except for the few cases explained in the footnote to Table 2a,* the number of licenses issued is identical to the number of years of teaching experience.

After assembling and tabulating the data for all the teachers, these significant facts and figures appeared:

Of the 265 cases, 191, or 72.08% were women; 74, or 27.92% were men. There were 145 elementary teachers, consisting of 124 women and 21 men; and 120 high school teachers, consisting of 67 women and 53 men. Even though the one factor held constant—that of having had at some time a license or

* See appendix for a sample of this application form, original thesis on file, Library, University of Wisconsin.

[1] This one exception has already been explained. He is the one case denied any other license after only one year of teaching. All other cases were allowed to complete at least the usual probationary period.

[2] For instance, Case No. 128 was granted a certificate after five years, and later received his Ph.D. degree and an instructorship in a large University. Other case studies are being prepared as a follow-up study of this thesis.

* All supplementary Tables, identified by letters (viz. 2a) are included in the appendix in the original thesis on file, Library, University of Wisconsin.

TABLE II

COMPARISON OF NUMBER OF LICENSES GRANTED WITH STATUS OF CANDIDATE

	Number of Licenses									
Status of Candidates	1	2	3	4	5	6	7	8	9	Total
Certificate granted	----	10	106	41	14	3	4	0	0	178
Certificate denied	1	10	22	1	2	1	----	----	1	38
Certificate pending	----	----	30	6	8	3	----	----	2	49

Ten of those who were granted a certificate had two licenses, 106 had three licenses, etc.
Chi-square = 7.77.
This is significant at the 5% level.
C = .17.

a certificate denied—the candidates grouped themselves into three classifications: that of eventually having (1) a certificate granted, or (2) a certificate denied, or of having (3) a certificate still pending (as of June, 1945). These three divisions are referred to as "status classifications", or simply "classification of candidates", and these terms refer to the professional status of a teacher as far as certification is concerned. (See Table III.)

Of these candidates 207 were graduated from the nine Wisconsin teachers colleges and Stout Institute, 32 from 19 of the county normals, 16 from the University of Wisconsin, and 10 from other colleges in this and other states. They went to teach during their probationary period in 497 different Wisconsin communities, ranging in size from open country to cities of over 10,000 population. But the majority of the communities, 439 of them, were rural as defined by the United States Census reports; that is they ranged in size from open country to towns of 2,500.

Much additional pertinent data regarding the cases selected for this study will be tabulated and discussed in the next chapter.

DEFINITION OF TERMS

It seems advisable at this point to insert definitions of some of the terms and concepts as they are used in this study. Any terms not included in this section are defined and clarified when they are first introduced.

1. *Probationary period.*—The first two years of teaching after being graduated from a recognized teachers' course given by a normal school or college, or after being admitted from another state. Two temporary licenses are granted, one each year, to the "probationers", unless supervisory officials recommend otherwise. This probationary period may be extended if a teacher has failed to "make good" during that time; or a person may be definitely disqualified for teaching in Wisconsin at the end of this two year period.

2. *Supervisory officials.*—Anyone acting in a supervisory capacity who visits a teacher and who, for this study, evaluates or rates the candidate. Those officials who acted in this capacity for the cases in this study include county and city superintendents, county, city, and state supervisors, and principals of schools.

3. *Candidate.*—A teacher on probation who is a candidate for a state certificate to teach.

4. *Criteria.*—Those aspects or characteristics or "qualities" of teaching which are used as the bases for the evaluation and the rating of the candidates in this study. These criteria have been selected by the State Department of Public Instruction and are listed on its recommendation blanks. The oral explanation[3] given state supervisors for the use of these criteria are:

1. Ability to instruct pupils. This includes the entire "teaching act" and its results, as far as they can be ascertained.
2. Ability to control pupils. This means not only "discipline" but control and management of the physical plant, materials, the room situation and so on.
3. Knowledge of subject matter. (No explanation is given. The inference is that it means "Does the teacher know the subject matter that he teaches?")
4. Professional interest and growth. This is determined by attendance at summer schools, institutes, and other professional meetings, keeping well read and in-

[3] Note that no explanation is printed on the reference blank or sent to any of the other people who do the rating.

formed on professional matters, and keeping "up to date" on procedures and techniques.

5. Cooperation. This merely means, "Does the teacher work well and generally 'get along' well, with other people?"

5. *Jurisdiction.*—The legal sphere of control and authority under which a school operates. The cases in this study all taught under the legal jurisdiction of one of two types of superintendents:

a. A county superintendent, who has authority over all schools in open country, towns, or suburbs which are not included in the district boundaries of a municipality; or

b. A city superintendent, who has authority over all schools within his municipal boundaries.

6. *Status* (of a candidate or teacher).

a. Certification status refers to the status of a candidate's certificate at the time of the study. His certificate has either been (1) granted, or (2) denied, or (3) is still pending. This grouping of certification status is also referred to as "classification of candidate."

b. Professional status has to do with one's professional reputation, as a teacher and professional worker. One's professional status can be effected by ones' certification status, and vice versa.

7. *Ultimate success* (of a candidate).—If a teacher, after one or more additional years of probation, finally is granted a certificate, he is considered an "ultimate success" as far as certification is concerned, and is put in the "certificate granted" classification.

8. *Ultimate failure* (of a candidate).—If a teacher is definitely denied either another one year license or a certificate (which means he can no longer teach in Wisconsin) he is considered an "ultimate failure" and is put in the "certificate denied" classification.

9. *Open-county schools.*—Not only one-teacher rural schools, but those of two or more teachers which are situated outside of towns or villages.

CHAPTER II

ANALYSIS AND INTERPRETATION OF THE DATA

Overview of the Chapter

While seeking answers to the questions which are pertinent to this study, and which are raised and discussed in this chapter, many interesting facts came to light regarding the teachers whose records are being examined. These findings seemed to group themselves about four factors which are involved in the professional life of the candidates, namely: (1) *The legal aspects* of certification, jurisdiction, and training; (2) *the institutions* in which these teachers were trained; (3) *the criteria* used for evaluation and recommendation purposes, and *the people who used these criteria;*[4] and, (4) *the situations* or *the communities* in which these candidates spent their probationary period.

These four factors must be seen in relation to each other; any one can be overemphasized if treated alone. There is a great deal of overlapping and interrelatedness in the data on each factor, and this must be remembered when reviewing any analysis and interpretation. The clusters of inferences which are derived from an analysis of these data, and which seem to group themselves about the four factors just mentioned, are treated in this way only for purposes of organization, clarification, and convenience, and not because they are in any way isolated from each other.

1. *The legal aspects.*—a. Licenses, certificates, and status of certification.—How many one-year state licenses are issued to these teachers who are having difficulty, before they are granted or definitely denied an unlimited certificate? How many such licenses have been issued to the current "pending" cases? In other words, how many chances do these candidates have, if they fail to receive a certificate at the end of two years? Does the number of licenses granted seem to make any difference with a candidate's chances of getting a certificate?

As has been previously explained, *licenses* are granted on a year to year basis; an unlimited or a five year non-renewable *certificate* in Wisconsin is granted when and if a

[4] This factor is also called the factor of "raters and rating criteria", for as a matter of fact; it appeared that most of the officials did actually "rate", rather than evaluate, the work of the teachers and used the criteria in that manner.

candidate has held two one-year licenses and has taught successfully for the concurrent years. A second license may be denied at the end of the second year, or subsequent years, of teaching.

An analysis of the data shows that the number of licenses held by these candidates did make a difference in their being granted a certificate; if one held from one to three licenses he was apt to be denied, and if he held from four to nine licenses he was apt to be granted a certificate.

In order to treat the data in Table 2 statistically, the columns were combined for "number of licenses granted", leaving the two columns of 1–3 and 4–9. (Table 2a)* Actually, only three of these candidates have been given 9 licenses, or nine years in which to "make good". One teacher was finally denied further licensing after nine chances (years); two teachers are in the "pending" classification, that is, they are now holding their ninth one-year license.

Four of those who had a certificate finally granted held seven licenses previously; three had six licenses, fourteen had five, and forty-one were certified after having four licenses. One hundred and six candidates were granted certificates after holding three licenses; that is, 106 of the 265 cases, or 40%, made good, as far as certification is concerned, in the third year of probation. (This is 59% of the 178 candidates who have been granted certificates.) Twenty-two candidates were denied certificates and further licensing at the end of three years of experience; that is, after extending the usual probationary period one more year. (This is 57.89% of the 38 cases who have been denied certificates.)

Does it make any difference to the future success or failure of these candidates whether they have already had from one to three or from four to nine licenses? Using the Chi-Square formula as a test, we find that a real difference does exist between the classification of candidates (whether certificates have been granted, denied, or pending) and the number of licenses held. The Chi-Square is 7.77 for Table II, and this is significant at the 5% level. That is, in all probability the chances are 95 to 100 that the number of one-year licenses held by a candidate who is having difficulty is a significant factor and is not due to chance alone.

* See Appendix in original thesis on file, Library, University of Wisconsin.

When those candidates who were finally successful in having a certificate granted are compared with those whose licenses are still pending, we find no significant difference between them. Any difference between these groups and number of licenses is due to chance. When, however, the classification of "granteds" are compared with the "denieds", and the "denieds" are compared with the "pendings", we find that there is a significant difference when the Chi-Square is used as a test. That is, any difference existing between these groups and the number of licenses held by them will be, 95 times out of 100, a real difference and one not due to chance.

Applying the Q[8] to each of the four-fold tables mentioned above, we find that the value of Q is +.56 when the "granteds" and the "denieds" are compared with number of licenses granted. This shows that there is a positive relationship between having 1, 2 or 3 licenses and having a certificate denied. That is, a candidate will probably be denied a certificate if he has from one to three licenses; at least, for these cases, there is more of a chance that they will be denied. From this we might also imply that if a candidate persisted and received a fourth, fifth or even ninth license, he will, probably, eventually be granted his certificate. (Tables 2b, 2c, 2d.)*

A common sense explanation of this seems to be that those candidates who were considered "hopeless cases", by the raters were denied further licenses or certificates at the end of the second or third year; and that those who were granted four or more licenses were given "the benefit of the doubt" and further opportunities to make good.

It has been explained that candidates have been classified into groups according to their *certification status;* that is, these cases fell into three categories; those who have had their (1) certificates granted, those who have had their (2) certificates denied, and those who have (3) certificates pending. Is there any difference between the men and the women in the elementary and high school fields when compared on the basis of this certification status? The data concerning this question were brought together in Table III and the Chi-Square formula was applied as a

[8] Yule's Q or coefficient of association. The Q portrays a real value if the x^2 has been previously shown significant for the fourfold table, thus there is no reason for computing the standard error of the Q.

* Appendix, original thesis on file, Library, University of Wisconsin.

TABLE III

COMPARISON OF WOMEN AND MEN IN ELEMENTARY AND HIGH SCHOOL FIELDS WHO HAVE CERTIFICATES GRANTED, DENIED OR PENDING

Classification of candidates	Elementary				High School				Totals				
	Men		Women		Men		Women		Men		Women		Total
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.
Certificate granted	14	67	80	65	42	79	42	63	56	76	122	64	178
Certificate denied	3	14	20	16	6	11	9	13	9	12	29	15	38
Certificate pending	4	19	24	19	5	10	16	24	9	12	40	21	49
Total	21	100	124	100	53	100	67	100	74	100	191	100	265

Chi-square = 5.55.
This is not significant at the 5% level.

test.[6] The value of 5.55 is not significant, therefore we must conclude that there is no significant difference between elementary and high school teachers, subclassified as men and women when compared with certification status. As long as there is no difference in the table as a whole, there is no difference in any segment of it, so that the statement just made also means that there is no significant difference between men and women as far as having a certificate granted, denied or pending is concerned.

It is just as important to know that there is *no* difference as it is to know that there *is* a difference. Knowing, then, that it makes no difference in one's chances of getting a certificate if one is either a man or a woman in the elementary or high school field, we must conclude that it is some factor or factors other than sex in relationship to field of service, which makes the difference in being denied or granted a certificate.

b. Jurisdiction of schools.—Did these teachers having difficulty with certification spend their probationary period in schools under city or county jurisdiction and supervision? In what proportion did they go to either place?

Did it make any difference with the chances these candidates had for ultimate success, if they were to teach in a school or schools that were under the jurisdiction of a county or a city superintendent, or if they spent the probationary period under *both* types of systems?

It was found that there were real differences, and that the candidates had a better chance for success if they stayed under the jurisdiction of *either* county or city jurisdiction and *not* under both.

A tabulation of the data shows that 211 or 79.62% of these 265 failures have been in county systems, 29 or 10.94% in city systems, and 25 or 9.44% have spent their first teaching years in both county and city schools. (Table IV) According to the Chi-Square value of this table wherein the candidates have been classified according to certification status, having, (1) certificate granted, (2) certificate denied, and (3) certificate pending, there is a statistical significance as far as the jurisdictional placement is concerned. This Chi-Square value of 11.40 is significant at the 5% level; which may mean that there are at least only five chances in 100 that the differences are due to chance.

When the ultimately successful teachers (certificate granted) were compared with those whose certification is still pending (classificated 1 with 3), and the Chi-Square formula was applied as a test, it was found there was no significant difference. The same held true when the ultimate failures (certificates denied) and the "pending" classifications (2 with 3) were compared. Any difference between these groups and jurisdictional placement is due to chance.

When, however, the successful and unsuccessful cases were compared (classification 1 and 2) it was found that there is a significant difference between those who finally were granted certificates and those who were finally, definitely denied, as far as the kind of system—whether county or city or both—was concerned.

Likewise, when the county jurisdiction data were compared with both (county and city)

[6] The Chi-square formula has been used throughout this study as a test of significant difference. The 5% level has been chosen as a basis for determining significance.

TABLE IV

ADMINISTRATIVE AND SUPERVISORY JURISDICTION OF SCHOOLS WHERE THE THREE CLASSIFICATIONS OF CANDIDATES SPENT PROBATIONARY PERIOD

Classification of Candidates	Number and % of Candidates who taught under the Jurisdiction of:							
	County		City		Both* County and City		Totals	
	No.	%	No.	%	No.	%	No.	%
Certificates granted	135	64	26	90	17	68	178	67.17
Certificates denied	36	17	1	3	1	4	38	14.34
Certificates pending	40	19	2	7	7	28	49	18.49
Totals	211	100	29	100	25	100	265	100.00

* One or more candidates in each classification spent the probationary period in 2 or more communities, some of which were under county and some city jurisdiction.
Total Chi-square = 11.40.
This value is significant at the 5% level.
C = .20.

data, and when city jurisdiction data were compared with both and county, the Chi-Square values show a significant difference between them. There is no difference between county and city, taken together, and "both". (Tables 4a, 4b, 4c, 4d, 4e, 4f.)*

It would seem, then, that it is not an "either-or" proposition as far as teaching in a city or a county system is concerned, but that the difference is a real one if a candidate teaches in both county and city—that is that he spends one or more years of his probationary period in a city system, and one or more years in a county system.

This may be interpreted to mean that those whose certificates are still pending would be more apt to have their certificate finally granted if they stayed in either a county system or a city system, and did not shift from one type of jurisdiction to another.

On the other hand an analysis of the data tabulated and treated statistically in Table XI, shows that approximately one third of the elementary teachers who are having difficulty with certification, have a better chance of finally acquiring a certificate if they go to more than one *community* during their extended probationary period. Nevertheless, these different communities should be under the administration of either a county superintendent or a city superintendent. In other words one third of the elementary teachers who are not succeeding might profitably teach in two or more different places within a county system, or within a city system, but should not shift from a county to a city, or vice versa, while having difficult with certification.

c. Training of candidates.—Did it seem to make any difference in a candidate's chances of success or failure if his training were for rural, elementary, or high school teaching?

Although the present certification regulations theoretically do not permit teachers to teach in fields other than the ones for which they trained, actually during the war emergency some teachers were given special permission to do so. Also, before the certification law took effect in 1939, it was not uncommon to find teachers who prepared for high school, going into graded or rural schools to teach and to find elementary teachers teaching grades for which they were not prepared.

The analysis given below points out that it did make a difference for these people and that most of the difference was in the rural field. Teachers in this field were more apt to be in the pending class than were the elementary or high school teachers.

The records of the cases selected for this study show that 223 taught only in the field for which they trained, but that 42 did not. Therefore, it seemed wise to find out whether or not there was any difference between teachers trained for rural, elementary, or high school, when compared with the status of their certificates. Table V was set up for this purpose, and an inspection of it shows that of the 145 elementary teachers, 62 trained for the rural-elementary field and 83 for graded elementary work. The 120 high school teachers all took high school training courses, (although nine of them taught elementary schools during their probationary period).

* Appendix, original thesis on file, Library, University of Wisconsin.

TABLE V

COMPARISON OF TYPE OF TRAINING WITH CLASSIFICATION OF CANDIDATE

Classification of Candidate	Trained for Rural Teaching	Trained for Elementary Grades	Trained for High School	Total
Certificate granted	31	63	84	178
Certificate denied	12	11	15	38
Certificate pending	19	9	21	49
Total	62	83	120	265

Chi-square = 12.85.
C = .21.
This is significant at the 5% level.

When the Chi-square is applied to Table V, we find a significant difference between the fields for which these people trained and their certification status. Only in five times out of 100 are the differences between these classification of candidates and their training, due to chance. The C. of .21 tells us there is a low but genuine relationship between those trained for rural, elementary, and high school, and the status of their certification.

When the elementary and high school trained teachers are compared, we find no significant difference; but when those trained for rural are compared with elementary and then with high school, we find that there is a real difference between them both. (See Tables 5a, 5b, 5c.)* There is then a real difference, and one due only rarely to chance, between those teachers trained for rural teaching and those trained for the elementary and high school fields. This means, that it does not make so much difference for having a certificate granted or denied if one trains for either elementary grades or high school teaching, but it does make a difference if one trains for rural teaching.

When those of certificate-granted status, are compared with those of certificate-denied, and those of denied status with pending, we find no difference between them. But when those candidates who have certificates granted are compared with those in the pending status, we find a value which is significant, and which shows that there is a real difference between those having a certificate granted and those having one pending, and the field of training. (Tables 5d, 5e, 5f.)*

Applying the Q to the fourfold table of certificate granted and pending, compared with rural and elementary training, we find a value of +.62. This tells us that there is a positive relationship between being trained for rural teaching and being in the "certificate pending" class: or that, for the cases under consideration here, one is more apt to be in the "pending" than in the "granted" group if one has been trained for rural instead of elementary teaching. However, when the Q is applied to a table comparing elementary with high school teachers, a positive relationship is found between elementary training and certificate pending. This means that those trained for elementary grade teaching have a better chance to have a certificate granted than do those trained for rural schools, and that those trained for high school have a better chance than do those who take elementary courses. (Tables 5g and 5h.)*

One might venture to explain this from several different points of view. Assuming that most teachers were placed in the type of school for which they were trained, it might be said that these rural teachers either (1) were not as well selected as the others, (2) did not get as good training, (3) did not get as much help during their probationary period, (4) had a more difficult job, or (5) that the people who evaluated their work were more discriminating or exacting in their recommendations than were those who evaluated the work of the others.

The same might be said when comparing these elementary teachers with those who were trained for high school. At any rate, it is obvious that no generalizations should be made regarding the reasons for these results, without further careful study.

2. *Institutions.*—What proportion of these candidates went to colleges, normals and the university? Does it seem to make any differ-

* Appendix, original thesis on file, Library, University of Wisconsin.

* Appendix, original thesis on file, Library, University of Wisconsin.

ence in the number and per cent of their below average ratings if they graduated from Wisconsin teachers colleges, county normals, the university, or from other colleges?[7] Is it possible to tell whether or not any one of these institutions does a better or a poorer job than any other in educating their students to be teachers, and in helping them develop ability to instruct and control, knowledge of subject matter, professional interest and growth, and cooperation?[8] Does any single institution do better in developing one characteristic than it does in another?

An analysis of data pertinent to these questions shows that there is no significant difference between these institutions and the performance of their graduates in the aspects of teaching just cited, according to the judgment of the people who rated them.

Data pertaining to this information are tabulated in Table VI. We find that 207, or 78% of the 265 cases were graduated from the nine Wisconsin teachers colleges and Stout Institute; 32, or 12% were graduated from county normals; 16 or 6% from Wisconsin University; and 10 or almost 4% from other colleges. These percentages have no meaning, of course, except a proportional one for the 265 cases in this study, and no generalizations can be drawn from these figures regarding the relative effectiveness of any of these schools.

In order to treat the data statistically, the "University" and "Other Colleges" columns were combined; but when the Chi-square formula was applied, it was found that its value was not significant. This means that 95 times out of 100 it made no real difference to these candidates, as far as their difficulties with these aspects of teaching were concerned, whether they graduated from teachers colleges, Stout, county normals, the university or other colleges. This does not say that all of these institutions do equally well, or equally poorly in the preparation of teachers; it simply means that for those cases who had difficulty there was no significant difference between the schools they attended and their rating on the five aspects of teaching accepted by the state for evaluation purposes. (Table 6a.)*

The (CR) formula was next applied to Table VI to ascertain whether or not there was any difference between the below average ratings of candidates in the different criteria. Only in "Knowledge of Subject Matter" was any significant difference established (CR = 1.96); all other criteria were found to have no difference between them. That is, any differences between ability to instruct, ability to control, professional interest and growth, and cooperation, in these candidates are due, 95 times in 100, to chance.

When comparisons were made in row 3 of Table VI (knowledge of subject matter) the only significant differences were between county normals and other colleges (2.34), and between teachers colleges with other colleges (2.71). Although it might be said that the people who rated these candidates felt that "other colleges" were preparing teachers in knowledge of subject matter better than the rest of the institutions, there are actually too few cases in the last column on which to base this statement. Because of the few cases, we are reluctant to accept this difference as being real, and feel that we should say that any difference between candidates from these various institutions in their knowledge of subject matter are, as in the other four criteria, probably due to chance.

However, when comparisons were made *within* each column, significant differences were found between criteria in three of the four columns.

(1) *Teachers colleges.*—There is a significant difference of 2.70 between knowledge of subject matter and cooperation. Therefore, there is a significant difference between knowledge of subject matter and all the other criteria, (because if a difference exists between the smallest and next smallest figures it will exist between the smallest and all the others). This means that the people who rated these 207 candidates from the teachers colleges and Stout, felt that these probationary teachers were better in knowledge of subject matter than in cooperation, than in professional interest, than in ability to instruct or control. In other words the raters feel that the graduates of teachers colleges know subject matter better than the other aspects of teaching.

[7] Stout Institute is included with the teachers colleges because it is primarily a teacher-training institution. For list of "Other Colleges" see appendix.

[8] These are the aspects or characteristics of teaching which are used as criteria for evaluation purposes by the State Department of Public Instruction. They will be discussed in more detail in a later section.

* Appendix, original thesis on file, Library, University of Wisconsin.

TABLE VI

CANDIDATES CLASSIFIED ACCORDING TO INSTITUTION FROM WHICH GRADUATED AND COMPARED ON THE BASES OF BELOW AVERAGE RATINGS IN THE FIVE CRITERIA

Marked below average in:	Teachers Colleges Including Stout			County Normals			University			Other Colleges			No. of cases rated below average
	No. of Grads.	No. Rated	% rated below average	No. of Grads.	No. Rated	% rated below average	No. of Grads.	No. Rated	% rated below average	No. of Grads.	No. Rated	% rated below average	
Ability to instruct	207	138	66	32	21	65	16	11	68	10	6	60	176
Ability to control	207	148	71	32	25	78	16	15	93	10	7	70	195
Knowledge of Subject Matter	207	79	38	32	13	40	16	6	37	10	1	10	99
Professional Interest and Growth	207	11	57	32	18	58	16	12	75	10	6	60	154
Cooperation	207	107	51	32	12	37	16	7	43	10	3	30	129
Total no. of cases given below average ratings* (cumulative)		590			89			51			23		753
	No.	%		No.	%		No.	%		No.	%		265
Total no. and % of cases	207	78		32	12		16	6		10	4		100%

* A candidate may appear once or as many as 5 times in these totals.

This table indicates the number of graduates from each type of teacher training institution; the number of candidates in each group who were marked below average in (1) Ability to Instruct, (2) Ability to Control, (3) Knowledge of Subject Matter, (4) Professional Interest and Growth, and (5) Cooperation; and the percentage marked below average. Of the 207 teachers' College graduates, 138, or 66%, were marked below average in ability to instruct, etc.

Chi-square = 3.81.

This is not significant at the 5% level.

This does not say that teachers colleges do a better job with subject matter preparation than all the other training institutions; it merely says that teachers colleges seem to have done a better job within themselves, and for these candidates, in the subject matter areas than in the so-called professional aspects of teaching; or that the teachers colleges do a poorer job in all the other four characteristics of teaching than they do in the knowledge of subject matter.

(2) *County normals.*—There were no significant differences until cooperation was compared with ability to instruct; then a difference of 2.33 was found. Therefore, there is also a significant difference between cooperation and ability to control. This means that when cooperation, knowledge of subject matter and professional interest and growth are compared, there is no difference; but when these *three* are compared with ability to instruct and ability to control a significant difference is found. This is to say that the raters of these 32 county normal graduates, felt that the latter did a better job in cooperating with others, in knowledge of subject matter, and in professional interest, taken together, than they did in manifesting an ability to teach or to control.

This does not mean that the evidence shows the county normals did a better job in these three aspects of teaching than did the other institutions; it merely means that within their own schools, the county normals obtained better results with these graduates in developing cooperation, knowledge of subject matter, and professional interest and growth, than they did in teaching these young teachers how to instruct and how to control. Or, stating it negatively, this means that the county normals did a poorer job in teaching these graduates how to teach (instruct) and how to control, than they did in the other three characteristics under consideration.

(3) *University.*—There was no significant differences found until knowledge of subject matter was compared with professional interest and growth. The significant difference between these criteria was 2.34. It follows then, that there is also a significant difference between knowledge of subject matter and ability to control. This means that these 16 graduates from the University did better, according to the judgment of the raters, in knowledge of subject matter, cooperation, and ability to instruct, taken together, than they did in ability to control and professional interest and growth. Or, that University graduates do not do as well in control or in professional interest and growth as they do in instructing, cooperating and knowing their subject matter.

Again, the evidence does not say that the University does a better or a worse job in developing these characteristics than any other institutions; it merely is making comparisons with these five criteria among its own graduates. For these candidates, the University did better in preparing them in subject matter, cooperation, and for instructing, then it did in professional interest and growth and controlling a school situation.

(4) *Other colleges.*—There are no significant differences between any of the five characteristics listed. The closest value was 1.86 between knowledge of subject matter and ability to control, but as this is not significant, and as there are too few cases in this column to deal with statistically, no statement can be made regarding comparisons of these values.

Apparently, then, the raters of these 265 teachers felt that the latter did as well in their respective positions whether they graduated from Wisconsin teachers colleges, Stout Institute, county normals, the University or other colleges. At least, any difference is due to chance, and not due to the fact of being a graduate of any certain institution. In other words, it can be said that *these* county normal graduates seemingly did as good a job for their purpose and their particular situation as did the graduates of any other colleges. It would be difficult to generalize about this statement, however, because we are here concerned with people who have been, or are, temporary or permanent failures. Only if the same evidence were found for successful teachers could a positive statement be made. Even then, consistency of raters, techniques, and interpretation of criteria, as well as other factors would have to be considered.

It should be noted again that "ability to control" was consistently ranked with aspects in which the graduates did the least well; and that "knowledge of subject matter" was always ranked with these characteristics in which these candidates rated the best.

3. *Raters and rating criteria.*—After a teacher has finished his first year of teaching,

and becomes a candidate for a second license, an application blank is sent to him from the Department of Public Instruction.* The applicant lists among other things, the names of his supervisory officials on this blank, and returns it to the certification office of the Department. The certification office then sends each person, who is listed as a reference, a recommendation blank on which to report his evaluation of the teacher's ability and progress. The same procedure is followed at the end of the second year of teaching, when the teacher becomes a candidate for a certificate; and this is continued each year until a candidate is either granted or denied a certificate.

An examination of the recommendation blank* shows the "qualities" on which each candidate is to be rated as above average, average, or below average. These "qualities" may also be called characteristics or aspects of teaching; in this paper they are likewise frequently referred to as "criteria", for in reality they are the five criteria which are to be used as a basis for evaluating or rating the work of a teacher. Local principals and superintendents, and state supervisors are sent the same kind of form; occasionally, also a county or city supervisor is requested to recommend a candidate.

Evidently, each "rater" is to use his own interpretation of the meaning of (1) ability to instruct pupils, (2) ability to control pupils, (3) knowledge of subject matter, (4) professional interest and growth, and (5) cooperation, for no explanation of the meaning of these terms is given.[9]

It is on the basis of one or more checks in the "below average" column that candidates are refused a license or certificate, or given a license when a certificate was requested. A classification and summary of the "further comments" (if and when any are given) is included in the appendix.

An analysis and interpretation of the data that are particularly pertinent to this section on "raters and rating criteria" follows.

a. The men and women who are rated.—Is there any real difference between men and women in their below average ratings on the five criteria that are used for evaluation purposes? To refine the question still further, is there any difference between men and women of different ages as far as these aspects or characteristics of teaching are concerned? The answer is that there are only chance differences between the men and the women and their below average ratings.

For these cases, 19 was the youngest age for beginning probationary teaching. The beginning age for the other age group was chosen arbitrarily, because 22 is the average age for graduation from a four-year college course. "Age" here means age at which candidates began teaching. The men and the women were then divided into four groups: men, 19–22 years old; men, 23 years old and over; women, 19–22 years old; women, 23 years old and over. (See Table VII.) For convenience, we shall call these groups "younger men", "older men", "younger women" and "older women". The numbers of below average marks given to the separate groups were tabulated, and it was found that the cumulative number of cases marked below average at least once was 753. The frequency distribution of the candidates' ratings, in order of the most frequent to the least, is as follows:

Ability to control	195
Ability to instruct	176
Professional interest and growth	154
Cooperation	129
Knowledge of subject matter	99
Total	753

When the Chi-square was used as a test, it was found that there was no significant difference between the men or women, in either age group, when compared with number of below average ratings. Therefore, it apparently makes no difference for those having difficulty, at least 95 times out of 100, whether one is a man or a woman, regarding one's rating on reference blanks. There really is, then, no significant difference between men and women as a whole in their below average ratings in the five aspects of teaching that are used in this study.

To test still further, however, and to see whether or not there is any difference between proportions of men and of women, p values were computed.

There are no significant differences between ratings for any of the five criteria for either

* Appendix, original thesis on file, Library, University of Wisconsin.

[9] These terms are frequently abbreviated in this study for the sake of convenience; for instance "ability to instruct pupils" is often merely written as "instruct" or "instruction".

TABLE VII

COMPARISON OF MEN AND WOMEN, DIVIDED INTO AGE GROUPS, WITH RATINGS ON CRITERIA

Marked below average in:	Men 19–22 years old Number	%	Men 23 and over Number	%	Women 19–22 years old Number	%	Women 23 and over Number	%	Total
1. Ability to instruct	18	24	31	41	98	51	29	15	176
2. Ability to control	19	26	35	47	107	56	34	18	195
3. Knowledge of subject matter	11	15	16	21	53	27	19	9	99
4. Professional interest and growth	17	23	27	36	83	43	27	14	154
5. Cooperation	12	16	21	28	70	36	26	13	129
Total	77		130		411		135		753

Chi-square = 1.22
This is not significant at the 5% level.

This table indicates the number and percentage in the age groups 19–22 and 23 and over, of both men and women who were marked below average in (1) Ability to Instruct, (2) Ability to Control, (3) Knowledge of Subject Matter, (4) Professional Interest and Growth, and (5) Cooperation. There were 74 men and 191 women, and the percentages are based on these figures. Seventeen men, or 23% of the total of 74 men, were marked below average in their ability to instruct, etc.

the younger men or the older women. This means that these two groups did not have any more difficulty with one aspect of teaching than they did with any of the others.

For the group of older men, however, we find a significant difference of 2.00 between knowledge of subject matter and ability to control, but only between these two criteria. This means that the older men were apt to have about equal difficulty with all aspects of teaching except control, and that they were apt to have more difficulty with control than all the others.

The younger women apparently had more difficulty with instruction and control than with the other three aspects, for the significant difference between knowledge of subject matter and ability to instruct was 3.00, and therefore we know there is a significant difference between knowledge and control also. This is to say that the difficulties with knowledge of subject matter, professional growth, and cooperation, experienced by this group, were approximately equal, but that their difficulties increased and they were marked below average the most often in instruction and control.

When differences between criteria were examined, (horizontal p values) it was found that in ability to instruct, there is a significant difference of 2.36 between the older women and the older men. Therefore there is a difference between the two groups of women, but not between the two groups of men. Apparently, then, these younger women and older men had more difficulty with instruction than did the older women and the younger men, according to the people who rated them.

There is a significant difference of 2.63 in ability to control, and of 2.00 in professional interest, between the older women and the older men; therefore the same interpretation holds for these two aspects as for instruction. That is, these younger women and older men had more difficulty in controlling a school, and in growing professionally, as well as in instructing, than did the older women and younger men.

When considering knowledge of subject matter and cooperation, a significant difference of 2.00 was found for the former, and of 2.87 for the latter, between the older women and younger women. Therefore there are no other significant difference between these groups for these factors.[10] This means, in effect that the older women had less difficulty with knowledge of subject matter and cooperation than did the younger women, but that there is no difference between the younger women and either group of men in these respects.

To summarize these data is to answer the specific questions raised at the beginning of this section:

1. There is no real difference between men and women, as a whole, in their below average ratings on the five criteria for evaluation.

[10] Because the only significant differences are found between the lowest and the highest percentages.

2. The differences between men and women when grouped in the four age classes, are these:

a. The younger men had about equal difficulty with all five aspects of teaching, but less difficulty in instruction, control, and professional growth than the older men and younger women.

b. The older men had more difficulty with control than with the other four aspects of teaching, and apparently had more difficulty with instruction, control, and professional growth, than did the older women and the younger men.

c. The younger women, within their group, had more difficulty with control and instruction than with the other three aspects; and seemed also to have more difficulty with instruction, control, and professional g r o w t h than did the younger men, and more difficulty with all aspects of teaching than did the older women.

d. The older women within their own group, had about equal difficulty with all five aspects of teaching, but less difficulty than all the other groups with all aspects, according to the judgment of the raters. In other words, of all these people who had trouble with meeting the requirements for certification, the women 23 years old and over, had less difficulty, and fewer below average marks, than all other three groups.

e. The two groups which had the most difficulty—the older men and younger women—experienced more trouble in cooperation and instruction than in any of the other aspects of teaching.

There seems to be no definite explanation for these findings, but one might hazard a few opinions regarding the causes for them. Maturity might be a factor to be considered, especially as far as the women were concerned. The older women had less difficulty in all respects than did the younger women or the men. Thus, it may be that for these women added maturity was an asset for those who had difficulty during the early years of teaching.

This does not explain why the older men had more difficulty, according to the raters of these candidates, than the younger men. Maturity does not seem to be a factor in this case. It may be that the younger men were more adaptable than the older ones and profited more quickly from experience and assistance. But this is not consistent with the fact that the younger women had the most difficulty of all the groups, unless one can prove that there is a difference in adaptability at this age.

Other causes that might be considered are that these younger men happened to go to schools that were better supervised, or that they were more carefully selected for their training and the communities to which they went. And as any differences between these men and women as a whole, when compared with all five criteria, are due to chance, it may be that many other factors than age need to be considered.

There may be several reasons why "ability to control" gave these teachers more below average marks than any other aspect of teaching. One is, that the majority of them may really lack the ability to control a school situation because of some personal, social, or mental lack. Another may be that they did not have enough help in this respect from their training institutions or, later, from their supervisors. A third reason may be that those who supervised their work—principals, supervisors and superintendents—may have different standards of "control" than did the faculties of the colleges or the teachers themselves. Thus, a principal might be judging control on the basis of what is termed "old fashioned discipline", while the teacher may have believed it to be desirable to allow children to speak and move about freely. And, finally, these supervisory officials might have considered control as of primary importance, or even given it undue importance, thus putting it first in their minds and marking against this criteria before and more than any others.

b. The numbers of persons rating candidates.—Is there any regulation regarding the number of people who must rate a candidate before he is denied a further license or a certificate? How many people passed judgment on these candidates before they were so denied? For these cases, did it seem to make any difference in the number of people rating them if they graduated from one college or another?

There apparently is no regulation regarding the number of recommendations needed for denial; for the records show that 32 of these

teachers were denied further licensing or certification on the basis of only one recommendation, 68 on the basis of two recommendations, 49 on the basis of 3, and 58 on the basis of four recommendations for denial. Only 58 candidates of the total of 265 had more than four different people recommend that they be denied a license or a certificate. Assuming that the larger the number of people who evaluate a candidate, the more reliable is the evaluation, one might question the advisability of denying these people further licenses on the basis of only one or two recommendations.

The records show, also, that there is no relationship between number of licenses applied for and granted, (which number also indicates number of years of probation allowed) and the number of persons likely to rate these candidates. That is, these cases were granted or denied a license or a certificate on the basis of recommendations from one to eight persons, regardless of how many times they made application for such a license or certificate. Case number 27, for instance, whose certificate is still pending after being given five licenses, has been rated by only two people; case number 28 has been rated by only two people, and his certificate is also pending after being granted his ninth license. Case number 2, after five licenses, was denied on the basis of recommendations from four people.

Now it should be remembered that reference blanks are sent to all people whose names are listed by the applicant on the application blank. Therefore the number of people who actually send blanks to the State Department depends upon a number of conditions:

1. Perhaps the candidate was visited by only one supervisory official, such as the county superintendent or his supervisor, during the year. In that case only his name would appear on the application blank. (Note that the county supervisor's name is not asked for, although he would ordinarily visit a teacher under county jurisdiction three or four times as often as the superintendent.) Naturally, if one name only is listed, one recommendation could be the only number received.
2. Even if from two to four or more names are listed on the reference blank, only one person may reply to the request for a rating of the candidate.
3. On the other hand, every person who is asked to do so may reply to requests for the return of the rating sheets.

Thus it is seen that the number of persons rating these candidates could be primarily due to the chance factors of visitation and replying to requests and therefore it is doubtful whether any inferences should be drawn even from a statistical treatment of the data. Of course, as has been said, what happens to these people is just as real as though no chance element were involved. Therefore, in order to test for any differences between them, these candidates were grouped as graduates of the various institutions where they received their pre-service education, and compared with the number of persons who rated them on the five criteria used for evaluation purposes.

Strangely enough, it was found that there was a real difference between these cases from various institutions and the number of raters. (Table VIII) To test still further, the teachers college cases were compared with the university and other college group, but the differences between these were found to be entirely due to chance. (Table 8a)* But when the county normal graduates were compared with the university and other college graduates, and again with the teachers college cases, a real difference was found between them. (Table 8b and 8c)*

When types of relationships were sought, none were found between the teachers college cases and those coming from the university and other colleges. But a positive relationship was found between the teachers college graduates and having from five to eight raters, and a negative relationship between county normal candidates and having from five to eight raters. (Tables 8d, 8e, and 8f)* This means that these graduates of teachers colleges, universities and other colleges were apt to have more people rate them for certification purposes than were the graduates of county normals, and that the latter were more apt to have from one to four raters than were all the others.

* Appendix, original thesis on file, Library, University of Wisconsin.

TABLE VIII

NUMBER OF PERSONS RATING GRADUATES FROM THE VARIOUS COLLEGES OR NORMALS

Number of persons rating candidates	Teachers' colleges	County normals	University and other colleges	Total
1	20	12	0	32
2	51	15	2	68
3	42	2	5	49
4	47	1	10	58
5-8	47	2	9	58
Total	207	32	26	265

Chi-square = 47.70
This is significant at the 5% level.
C = .38
20 graduates of Teachers Colleges and 12 graduates of county normals were rated by only one person, etc.

This might be explained by the fact that county normal graduates almost always go to one or two room schools to begin their teaching, and that these schools are visited only by the county superintendent and his supervisor. The graduates of other institutions, who usually have taken a three or four year course, are apt to go to small towns and villages, where they will be visited by the principal of the school, the county superintendent and supervisor, and a state supervisor. Thus the latter would have more people to list as references and would have a better chance of having more blanks returned.

Perhaps, from this point of view, the number of raters did not depend so much on chance as on the situation to which these teachers went. At any rate, the chances for having more than four people rate them, were against these graduates of county normals.

c. Types of persons rating candidates.—Are there any differences between the criteria used for evaluation purposes, when compared with the people who rate these candidates? Is there any consistency among raters, or any relationship among criteria?

The evidence shows that there are real differences, due only five times out of 100 to chance, when the raters and the rating criteria are compared. The findings also seem to indicate that there is a lack of consistency among raters and rating criteria, and that knowledge of subject matter is the only aspect on which there is some agreement among the evaluators.

It has been pointed out that the people who sent reference blanks to the State Department regarding their evaluation of the work of these candidates, included principals, county and city superintendents, and county, city, and state supervisors. (No one acquainted with the work or worth of the candidate in his pre-service education was consulted.) When the raters were compared with the rating criteria, a difference not due to chance was found. (Table IX) When further tests were desired and applied, in order to find the differences, if any, between various combinations of raters, it was discovered that there were no significant differences between any groups or combination of groups except between principals and all other raters combined. (Tables 9a, 9b, 9c, 9d, 9e, 9f)* There was a real difference between the ratings on these five criteria given by the principals when the latter were compared with all the rest of the raters.

Relationships or correlations were found, however, only when (1) superintendents (county and city), and principals, were compared with ability to instruct and ability to control; and when (2) principals and county superintendents were compared with ability to instruct and control. (Tables 9g and 9h)* These correlations were found to be negative between the superintendents and control, and positive between the principals and control.

Endeavoring to test still further for consistency of raters and rating criteria, the C. R. was applied to the data (tabulated in Table IX and in supplementary tables) in order to find vertical and horizontal p values. The only significant difference found in the former analysis was between knowledge of subject

* Appendix, original thesis on file, Library, University of Wisconsin.

TABLE IX

NUMBER OF CANDIDATES MARKED BELOW AVERAGE FROM 1 TO 4 TIMES BY DIFFERENT TYPES OF SUPERVISORY OFFICIALS

Candidates	Number of Candidates Given Varoius Ratings by: Principal				County Superintendent				County or City Supervisors				State Supervisor				City Superintendent				Total
Marked below average in:	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	
Ability to Instruct	60	4	0	0	71	18	2	1	16	4	1	0	35	12	2	0	10	1	0	0	237
Ability to Control	91	11	2	0	64	12	5	1	13	3	1	0	27	10	1	0	16	5	0	0	267
Knowledge of Subject Matter	25	1	0	0	34	9	3	0	7	2	1	0	23	4	1	0	5	1	0	0	116
Professional Interest and Growth	55	7	0	0	51	6	3	1	11	2	1	0	32	6	1	0	12	2	0	0	190
Cooperation	50	10	0	0	36	9	2	0	8	1	0	0	23	3	1	0	11	5	0	1	169
Cumulative total number of candidates		325				328				76				181				69			979

Chi-square = 31.09.
This is significant at the 5% level.
C = .17.
This table shows how many candidates were marked below average in (1) Ability to Instruct, (2) Ability to Control, (3) Knowledge of Subject Matter, (4) Professional Interest and Growth, and (5) Cooperation, by a Principal once, twice, etc., County Superintendent once, twice, etc.

matter and ability to control when rated by the principals. A summary of the horizontal p values for Table IX is included in the appendix.

What do all these analyses tell us regarding the results of the rating of these candidates by the people who use the rating criteria? Among other possible interpretations, the findings may be summarized by stating these:

1. This arrangement of these data (Table IX) seems actually to be a test of the ability of these raters to rate, and of the criteria used for rating.
2. The more people who rated a candidate, the more agreement there was on the particular criterion. There was less agreement on instruction than on any other criterion, and more on knowledge of subject.
3. Principals seemed to be more discriminatory than any other raters, but they r e a l l y discriminated only between knowledge of subject matter and control. The principals were also more consistent in rating instruction than any other supervisory officials. Principals agreed with the county superintendents in all aspects except cooperation.
4. County and city superintendents were evidently not putting the same emphasis on these factors. They did not agree on any counts, but agreed less on instruction and control than on the others. In fact, the city superintendents seemed inconsistent on all their ratings.
5. State supervisors agreed with principals only when rating subject matter but they agreed with county and city supervisors on all criteria and on all counts except instruction.

Probably many explanations could be found for this inconsistency within the criteria and lack of agreement among raters; but only a few are ventured here.

1. Different interpretations and meanings of the concept of "control" have already been discussed in another section. There seems to be little doubt but that this is a vital factor in the difficulty these candidates have. This factor of control should probably be singled out for special study and consideration because it comes first in difficulty for all these "failure cases". Different interpretations

TABLE X

COMPARISON OF THE CLASSIFICATIONS OF CANDIDATES WITH NUMBER AND PER CENT OF BELOW AVERAGE RATINGS

Below Average in:	Certificate Granted						Certificate Denied						Certificate Pending						Number of Cases Rated Below Average (Cumulative)
	Elementary			High School			Elementary			High School			Elementary			High School			
	Number of Candidates	Number of Candidates Rated	% of Candidates Rated	Number of Candidates	Number of Candidates Rated	% of Candidates Rated	Number of Candidates	Number of Candidates Rated	% of Candidates Rated	Number of Candidates	Number of Candidates Rated	% of Candidates Rated	Number of Candidates	Number of Candidates Rated	% of Candidates Rated	Number of Candidates	Number of Candidates Rated	% of Candidates Rated	
Ability to Instruct	94	74	78	84	43	51	23	18	78	15	11	73	28	17	61	21	13	61	176
Ability to Control	63	67		67	79		18	78		15	100		16	57		16	76		195
Knowledge of Subject Matter	39	40		29	34		6	26		5	33		14	50		6	29		99
Professional Interest and Growth	56	59		46	54		14	60		10	66		14	50		12	57		154
Cooperation	52	55		42	50		9	39		9	60		9	32		8	38		129
Number of Cumulative Cases	284			227			65			50			70			57			753
Total number of cases	94			84			23			15			28			21			265

Chi-square = 13.85.

This is not significant at the 5% level.

This table shows number of candidates who were granted a license, denied a license, or whose case is still pending, the number of these candidates who were rated below average in the particular trait, and the per cent of candidates rated below average. There were 94 elementary teachers who were granted a license, 74 of them or 78% were rated below average in ability to instruct.

of all these criteria, in fact, by many separate raters, is no doubt a major cause of the inconsistency in rating.

2. A lack of understanding on the part of the candidates regarding evaluatory and rating procedures is another explanation. Many teachers do not know what their supervisors "expect" when these five aspects of teaching are discussed.
3. Principals are closer to the teacher and the school situation than are any other officials. This may explain the fact that they are at the same time more consistent and more discriminatory than the other raters.

Further discussion of these findings, and their implications for the educational programs, will be delayed until the next chapter.

d. General comparison of criteria with certification status.—When the three classifications of candidates—those having (1) certificates granted, (2) certificates denied, and (3) certificates pending—are compared with the below average ratings on the five criteria used for evaluation purposes, is any significant difference found between these groups and their number of ratings?

At first glance, it may seem strange to say that no significant difference was found to exist between these factors when the data were treated statistically. (Table X) But the Chi-square is a test of independence as well as of significant difference;[11] so the fact that these data do not prove to be significantly different, does prove dependency between factors.

This is what would naturally be expected from the analyses of this study. For these candidates became "cases", and fell into these categories, because they were marked below average on one or more criteria at least once. Therefore the factors brought together in Table X are mutually dependent upon each other; and any differences which do occur between certification status, (subdivided into elementary and high school groups) and below average ratings are due to chance.

This means that, generally speaking, there were no differences between these elementary and high school teachers, in any of the classifications, when compared with the number of below average ratings on the criteria. If any differences did exist between these groups, then, other factors than ratings on criteria must have caused them.

4. *Situations or communities where probationary period was spent.*—a. Number of communities.—Do these candidates usually stay in one community for their probationary teaching, or do they move about? Are there any differences between the elementary and High School teachers in this respect? Are there any differences between classifications of candidates (those who have had certificates granted or denied, and those whose certificates are still pending) regarding the number of communities in which they teach during the probationary period?

The evidence seems to indicate that there is a real difference, and that the relationship between elementary teachers and number of communities is significant and positive. For high school teachers, however, it seemed to make no difference whether they went to one or more than one situation.

Table XI shows that 96 of the 265 teachers having difficulty with certification taught in only one community during their probationary period, 118 taught in two communities and the remainder of the 265 spent their first years in from three to five communities. Of this latter group, 40 taught in three different places, ten taught in four, and one candidate taught in five. All together, these 265 candidates spent their probationary period in 497 different communties. (See also Table XII)

When the Chi-square is used as a test of significant difference for Table XI, we find its value of 19.19 to be significant. This means that at least 95 times out of 100, there was a real difference between these elementary and high-school teachers and the number of communities where they spent their probationary period; or that in only five times out of 100 would a difference between them be due to chance alone. The C value of .26, tells us that there is a low but genuine relationship between these elementary and high school teachers and the number of communities in which they taught during this period.

After finding that Table XI was significant, a comparison of the three classifications of high school teachers alone was made with number of communities, only to find that no significant difference exists; that is, it made no real difference for these high school teachers, as to whether they had their certificate granted, denied, or still pending, in how many

[11] R. A. Fisher, *Statistical Methods for Research Workers* (London: Oliver and Boyd, 1938), p. 88.

TABLE XI

NUMBERS OF CANDIDATES IN THE THREE CLASSIFICATIONS THAT SPENT PROBATIONARY PERIOD IN 1, 2, AND 3–5 COMMUNITIES

Classification of Candidates	Elementary Teachers			High School Teachers			Total
	1 Community	2 Communities	3–5 Communities	1 Community	2 Communities	3–5 Communities	
Certificate granted	29	46	18	32	33	20	178
Certificate denied	14	8	1	4	10	1	38
Certificate pending	10	11	8	7	10	3	49
Total	53	65	27	43	53	24	265
		Elementary		53	65	27	
Combined Total, Elementary and High School				96	118	51	265

Twenty-nine elementary candidates who were granted a certificate spent their probationary period in only one community, 46, in two different communities, 18 in from 3 to 5 communities, etc.
Total chi-square = 19.19.
This is significant at the 5% level.
C = .26.

different places they taught during probation. Likewise when the three classifications of both elementary and high school teachers were compared with all combinations of numbers of communities, and when the "granted" and "denieds" were compared with one and more than one community and with elementary and high school teachers, no significant difference was found. (Table 11a, 11c, 11d, 11e, 11f, 11g)*

However, when the classifications of elementary teachers alone were compared with number of communities, a close but significant value was found. (Table 11b)* Knowing, then, that the difference between classification of elementary teachers (having certificate granted, denied or pending) and teaching during the probationary period in one or more than one community, is real, our next question is; what does this mean for these elementary teachers? ?

When we compare elementary teachers who have had their certificates granted or denied, with one community, and more than one, we find a c of .38 and a Q of +.55. (Table 11h)*

This means that for these elementary teachers there was a positive relationship between spending their probationary period in more than one community and being granted a certificate. It also means that there was a positive relationship between spending that period in one community and being denied a certificate. In fact, approximately 30 times out of 100, those elementary teachers who were having difficulty, were more apt to have a certificate granted if they taught in more than one community, than they were if they taught in only one.

This held true only for the elementary teachers, however, and not for the high school. There seemed to be no relationship whatever between spending the probationary period in one or more communities and being granted or denied a certificate for high school teachers.

We can state, however, that at least almost one third of the elementary teachers who have had difficulty with certification will be granted a certificate if they go to more than one community during their probation, or will be denied if they stay in one community. This does not say that the other 70% will be granted or denied on the basis of number of communities in which they teach. Too many other factors enter into the success or failure of teachers to make any such statement. But for any other similar group of elementary teachers, it would be best, in 30 times out of 100, to teach, during probation, in more than one community.

This is, of course, only a good, commonsense procedure, and would be recommended by any alert supervisor. If a teacher is in difficulty in one community, and has any possibilities of "making good", he should be transferred to a different community and given another chance. For these high-school teachers however, other factors seemed to have been involved.

* Appendix, original thesis on file, Library, University of Wisconsin.

b. Size of communities.—Did more of these candidates go to small or to large communities for their first years of teaching? In what proportion? To how many communities of less than 2500 or of more than 2500 population, did the candidates who have certificates granted, denied, and pending go? Is there any significant difference between the three classifications of cases and the size of the community? Is there any significant difference between elementary and high school teachers compared to size of community?

Most of these teachers went to small communities of 2500 or less population for their probationary period. There was a real difference between the high-school and the elementary teachers in this respect, and a relationship between size of community and status of certificate if one were an elementary teacher.

An inspection of Table XII and 12a* will give us the answers to more of these questions. Altogether, these 265 cases went to 497 communities for their probationary teaching. Of these 497 places, 439 ranged in size from open country to 2500 population, and only 58 were in the 2500 and over class. This means that 88% of the communities included in this study would be classified by the United States Census as "rural", and that 12% would be classified as "urban". The 178 candidates who had certificates granted to them after an extended probationary period went to 284 rural communities and 50 urban ones; the 38 who were denied certificates went to 58 rural and two urban communities; and the 65 whose certificates are pending have gone so far to 97 rural and six urban centers.

It is evident then, that the first difficult years of these teachers were spent largely in rural and village environments; that is, in places small enough to allow teachers to take a real part in community living, and to give community members an opportunity to become acquainted with teachers as people as well as professional workers. Does this make any difference in the chances these teachers have for success or failure?

Columns were combined as shown in Table 12a* so that the table could be treated statistically. When the Chi-square formula was applied, a value that was significant was found. This indicates that 95 times out of 100 the differences between the status of candi-

* Appendix, original thesis on file, Library, University of Wisconsin.

TABLE XII

NUMBER OF COMMUNITIES OF DIFFERENT SIZES WHERE THE THREE CLASSIFICATIONS OF CANDIDATES SPENT PROBATIONARY PERIOD

Classification of Candidates	Number of Communities Elementary Teachers						Number of Communities High School Teachers						Total number of communities	Total number of teachers
	Open country	Up to 250	250–1000 to 1750	1000 to 2500	1750 to 10,000	Over 2500	Open country	Up to 250	250–1000 to 1750	1000 to 1750	1750 to 2 00	2500 to 10,000	Over 10,000		
Certificate granted	92	9	28	11	4	9	20	20	28	69	18	5	10	334	178
Certificate denied	21	3	6	3	2	0	0	2	4	14	3	0	1	60	38
Certificate pending	41	4	11	8	1	0	1	3	3	18	6	2	3	103	49
Total		16	45	22	7	9	21	25	35	101	27	7	14	497	265

15.4

Chi-square = 22.74.
This is significant at the 5% level.
C = .21.

dates, when compared with the size of communities to which elementary and high school teachers went, were real, and not due to chance.

In order to find out *within* the table where the differences lie, p values were found for the vertical and horizontal sections of the table. To facilitate the discussion at this point, the communities where the elementary teachers went will be called "elementary-communities", and those where high-school teachers went, "high-school-communities". When, then, the size of the community is held constant for (1) open country (2) 250 to 2500, and (3) 2500 and over and the status of the candidates varies, we find that the difference is between granted and pending (C. R. = 2.25) and granted and denied (C. R. = 4.80), in elementary-communities in the open country and elementary-towns of over 2500. There is no differences whatsoever for any status classification in any high-school-communities.

When the status of the candidates is held constant, (for (a) certificate granted, (b) certificate denied, and (c) certificate pending), and the size of community varies, we find significant differences between (a) open country, elementary and high school (C. R. = 7.71), and 250–2500, elementary and high school, (C. R. = 6.18) for the "granteds"; (b) open country, elementary and high school (C. R. = 3.33) for the "denieds"; and (c) open country, elementary and high school (C. R. = 5.19) for the "pendings".

The open country was the actual significant difference all the way through, and this was a real difference between any of the classifications of teachers.

To test for correlations between certification status and elementary-community size and high-school-community size, three fourfold tables were derived. (See Table 12b, 12c, 12d)* No difference was found between the "denieds" and the "pendings", or between the "granteds" and "denieds". A significant difference was found, however, between the "granteds" and the "pendings", so the Q formula was applied. A positive correlation between the pending classification and the combined number of communities where elementary teachers go (elementary-communities) was found. This means that there is a negative correlation between having one's certificate granted and elementary-communities, and that there is less chance of getting a certificate granted if one goes to the elementary-communities than if one goes to the high-school-communities. In other words, it really implies that in terms of size of community there is less chance of getting a certificate if one is an elementary teacher. For high school teachers, it seems to make no difference for having a certificate granted, whether they go to smaller or larger communities. The real relationship between size of community and status of candidate, is contributed by elementary teachers who go to the open country to teach.

c. Relation of size of community to home community.—Did it make any difference with the chances these candidates had, if they went to a community larger, smaller, or of the same size as their home community for their first years of teaching? What proportion of these teachers went to their home communities, community of similar size, or larger or smaller towns?

The records include the name of the city or village in which the candidate graduated from high school, as well as the names of the communities in which he spent his probationary period. By considering the town where the candidate graduated from high school as his home community, the relationship of the size of the community in which he spent his first years to that of the size of his home community can be studied. Applying the Chi-square formula to Table XIII as a test, we found that it makes no difference whether those teachers who had difficulty with their certification went to their home communities to begin teaching, or went to a community of the same size, larger size, or smaller size. (Table XIII) It would seem, then, that size of community in relation to home community was only a chance element in the failure of these teachers. Even though the data show that a larger proportion of these failures went to communities *smaller* than their home towns (48%), than went to communities *larger* than their home towns (9.8%), for their first years of teaching, it is not a significant factor in their failure.

This evidence is not consistent with the theory generally held that beginning teachers have a more difficult time if they go to their home communities to teach; or that they have more difficulty in making adjustments to communities which are smaller than their home

* Appendix, original thesis on file, Library, University of Wisconsin.

TABLE XIII

RELATIONSHIP OF SIZE OF COMMUNITY IN WHICH CANDIDATES TAUGHT TO SIZE OF HOME COMMUNITY, COMPARED WITH CLASSIFICATION OF CANDIDATES

Type of Community in relation to size of home community	Classification of Candidates: Certificate granted	Certificate denied	Certificate pending	Total %	Frequency Totals
Home Community*	32	10	11	20.00	53
Same size as home community	41	9	8	21.89	58
Larger than home community	22	2	2	9.81	26
Smaller than home community	83	17	28	48.30	128
Totals	178	38	49	100.00	265

* Home community here means community in which candidate graduated from high school.
Chi-square = 6.56.
This is not significant at the 5% level.
Row 1. 32 candidates who were eventually granted certificates, 10 candidates who were denied, and 11 whose certificates are pending went to their home communities to teach. This is 20% or 53 of the total of 265 teachers.

communities. This study is dealing only with teachers who are already having trouble with certification, however; so perhaps other factors surmount the adjustment to home community or to towns that are of different size than the home community.

That *size* of community is of itself a significant factor has already been pointed out; but this does not mean size in *relation to that of home towns*. For these teachers, it did not seem to make much difference in that respect.

d. Communities with largest frequency of cases.—The fact that the situation in which beginning teachers find themselves may be a significant factor in their success or failure, should be examined further. An analysis of the data shows that 14 Wisconsin communities each had three or more of these candidates in the last ten years. The population of these 14 villages ranges from 315 to 2,621 according to the 1942 Blue Book. Nine of the villages had three candidates each, four communities had four each, and one small town of 794 population had five beginning teachers who were not recommended for a certificate at the end of the probationary period of two years. (Table XIV)

Four of the candidates in this study taught in two of these 14 communities where, altogether, 44 beginners had difficulty with certification. Case numbers 4 and 114 were placed in Community B, where they and two other beginners had difficulty with certification recommendations. Case Number 4 also went to Community K and Case Number 114 to Community I. Case Number 69 spent some of her probationary period in Communities D and E, each of which had three other "failing" teachers. Case Number 1 went to both Community E and Community H, the latter having had also two other beginners who had difficulty during the probationary period.

All four of these candidates whose records appear twice in the most frequently cited towns of this study, have been finally granted certificates after four years of probation—two years over and above the usual period. A question might very properly be raised as to whether or not these two men and two women could have begun their teaching careers as "successes" rather than "failures" if they had been placed in communities where beginning teachers seem to have a better chance of succeeding.[12]

Nine people who have had professional and/or social contacts with communities A, B, C, D, and E—those communities which have had the greatest number of beginning failures—have been interviewed. Among these were former residents, principals, and superintendents, and state supervisors. A summary of their statements regarding the school situation in these communities is given below.

Community A.—"This town refuses to pay enough money to hire or to keep good teachers. The school board always looks for inexperienced teachers, thinking and hoping to get them more cheaply. The principals who go here have as much difficulty as the teachers, and the board ignores them as far as the hiring of teachers is concerned. Only teachers who can't get jobs any place else will go to

[12] Further and more intensive study is already being made, by the writer, of these 4 cases, the 5 cases who went to Community A, and of Communities A, B, D, E, H, I, and K. The intention is to watch the history of these communities over another 5 year period and study the cases of all beginning teachers going here during this time.

TABLE XIV

NUMBER AND CLASSIFICATION OF CANDIDATES PLACED IN THE FOURTEEN MOST FREQUENTLY CITED TOWNS

Case No. of candidates who were placed in more than 1 of these com-	No. of candidates placed in these communities who have certificates:	Total no. of candidates

TABLE XIV

NUMBER AND CLASSIFICATION OF CANDIDATES PLACED IN THE FOURTEEN MOST FREQUENTLY CITED TOWNS

Community	Population	Types of Jurisdiction	Case No. of candidates who were placed in more than 1 of these communities and counted twice in this table	No. of candidates placed in these communities who have certificates: Granted	Denied	Pending	Total no. of candidates
A	794	County	----------	5	0	0	5
B	456	County	4 and 114	3	0	1	4
C	1116	City	----------	4	0	0	4
D	2621	City	----------	2	0	2	4
E	979	County	69	3	0	1	4
F	315	County	1 and 69	3	0	0	3
G	2382	City	----------	2	0	1	3
H	1034	City	----------	2	1	0	3
I	902	County	1	3	0	0	3
J	1362	County	114	1	1	1	3
K	584	County	----------	1	0	2	3
L	317	County	4	0	2	1	3
M	879	County	----------	3	0	0	3
N	1762	City	----------	3	0	0	3
Totals (14)		5 City	4*	35	4	9	48—4=44*
		9 County					
		14					

* The number of persons affected is actually 44, because 4 candidates of the total 48 appear twice.

this community; if they know anything about the town they won't teach there. After hiring the cheapest, most inexperienced teachers they can find, the board refuses to be patient with them or help them. The social level is low, and living conditions very poor."

Community B.—"The principal seems to be at fault here; the same man was in charge of the school when three of the four candidates were teaching here. He is not helpful to beginners; on the contrary he is quite fault-finding." (Although this principal did mark three of the candidates below average, a state supervisor marked two of them similarly.)

Community C.—"The school board and the people of this community expect perfection from all their teachers. The teachers' business is everybody's business, and teachers are closely watched in and out of school. People are critical of teachers and teachers feel insecure. There is a former teacher on the board, and although she was a good teacher she is a 'terrible' board-member!"

Community D.—"The board and superintendent do not hire enough teachers. As a result there is a big teacher-load and teachers often have to teach 'out of their own field.' This is naturally hard on beginners."

Community E.—"This is a Polish-Catholic community where there is a very heavy teacher load. Living conditions are poor and the social level is low. The principal hires and fires and does not help his teachers much. Although he seems to have an interest in children, he lets his teachers either 'sink or swim' without his help."

With these facts, and with these opinions of reliable people before us, might we not pose some such questions as these for further investigation?

Do some beginners already have "two strikes against them" if they go to a community where four or more, or even three or more beginning teachers have already failed to make good during the normal probationary period?

Could some communities be called more "dangerous" than others for beginners?

Could these *communities* be "below average for beginners" rather than the teachers being below average in the classroom? Could poor living conditions, heavy teaching load, critical watching of teachers, *cause* a beginner to do below average teaching?

Should our placement services study community facilities, and possibilities for beginning teachers within a given community before placing candidates therein?

Does a community have any responsibility toward teachers, especially beginners, or is the responsibility for making good entirely the teacher's?

Of course, there are many factors and variables other than community types, mores, expectancies and so on. Some teachers may never become successful, even in the most advantageous situation. Nevertheless it seems likely that one would question the advisability of sending beginning teachers to communities which, as the facts tell us, have had four or five, or even, possibly, three beginning failures in a ten year period. And one might be inclined to conclude, that the four teachers who went to two such communities were unfortunate indeed in their situations for the probationary period.

CHAPTER III

SUMMARY AND RECOMMENDATIONS

The main purpose of this study was to derive, from an examination, analysis and study of the state records of teachers who have had difficulty with certification, recommendations for the improvement of teacher certification and education in Wisconsin. Before such recommendations are made, however, it would seem best to summarize first some of the findings that have emerged from the examination and study of the source materials and from the investigation as a whole. The summary statements will be organized as replies to questions; and in turn, these answers and the implications drawn from them, will have reference to the problems raised at the beginning of this report. From these concluding statements and their implications, recommendations will be derived and formulated.

Summary statements and implications:

I. What are the sources, causes and results of the failure of some teachers to have certificates granted to them at the end of the usual probationary period, and what are their implications for teacher education and certification?

Undoubtedly, the sources of failure for some of these cases, lie in the personal, mental, social and emotional inadequacies of the teachers themselves. But if the question could be dismissed with this statement, it would imply that grave criticism would be directed against the teacher-education institutions. For the theory of careful selection of teacher-candidates is one that is commonly accepted, and most colleges claim to be practising it. If they are actually doing so, then the fact of personal unfitness cannot be accepted as an important source of the trouble.

Granting that errors of judgment in selection can, do, and probably will continue to exist, however, the factor of personal lack can still not be cited as a reason for so many of these failures.[18] For if the candidates passed the initial selection tests, they continued in that institution and were (1) allowed to pursue the courses prescribed for teachers (2) graduated with the major in education, and (3) recommended for a license to the State Superintendent. Therefore, any attempt to dismiss these people as being potential failures, only implies poor policies and practices of selection, education, guidance, graduation, and recommendation on the part of the college.

Other reasons for not accepting personal inadequacies as an important source of difficulty have already been pointed out; for instance, 67% of these cases did finally acquire a certificate, and several have become outstandingly successful in the teaching profession. If, then, this factor cannot be called any but a negligible source of difficulty, what are some acceptable sources and causes of these failures?

a. Some of the causes of difficulty lie in the problems these teachers had as they were beginning their teaching careers. First, the problems of greatest concern to most of them as reported by the supervisors, were the classroom problems of (1) controlling a school situation and (2) instructing the pupils in their care. The professional problem of showing evidence of (3) interest and growth in one's work, and the personal one of (4) cooperation with others, come next in line. Having a sufficient (5) knowledge of subject matter was a problem, but this was not nearly so important as the first four, and would probably not have caused trouble for these

[18] It should be remembered that the cases selected for this investigation do not constitute all cases of difficulty.

people if their other difficulties had not been so apparent.

All this might imply that these candidates could learn from books, for most of them, as has just been said, had little difficulty with knowledge of subject matter. But present theories of education contend that "controlling" and "instructing" in a classroom are as much a matter of *practice* as of "learning" in the abstract. If this is true, and if, in general, these people were as "fit to teach" as were those who were successful, this implies among other things, that (1) they did not acquire an understanding of the meaning or the theory of control and instruction in college; (2) they did not get enough practice in these aspects of teaching before they graduated; (3) they were put into situations for which they were not prepared; (4) their supervisors or the people in the community had different theories regarding these phases of teaching, and their performance was judged by a different standard than their own; or (5) that they were not given enough and proper help during their probation period.

The fact that the older women of this study had less difficulty with these problems (as well as with the others), than did the rest of the teachers, implies that added maturity is a factor for success, as far as women teachers are concerned. This coincides with the prevailing theory and practice of lengthening the preparatory courses for teachers, so that they will not only learn more, but also be more mature when they become teachers.

Other problems that these beginning teachers had were problems of placement; that is, of finding the "right" situation. These teachers had more difficulty if they moved from one type of jurisdiction to another; and the communities to which elementary teachers went were more of a factor in their problems than were those for the high school teachers. This was especially true for elementary teachers who were trained for rural teaching and who went to the open country to teach. This implies among other things, that (1) it is as important to study the communities for the teachers, as it is to find the right teacher for the community; and that (2) there are factors other than merely "teaching school" which affect a candidate's success or failure. It also implies either that (3) these rural and elementary schools, and the communities in which they are situated, were the most difficult in which to teach, or that (4) people who were less mature, and less well trained, went into these communities; *or* that (5) the supervisory officials who judged and reported on the work of these elementary teachers were more discriminating than were those who rated the high school teachers. At any rate, if the third and fourth implications are true (and most educators would agree that they are) and teachers are legally permitted to go to these difficult situations, it might be implied that more assistance and time should be given them before they are denied further licensing or certification.[14] From this one might infer that a longer probationary period should be allowed for those teachers who have had less training than the others.

These are not the only problems which confronted the teachers in this study, but they constitute some of the major ones that were reported by supervisory officials. As has been pointed out before, the "interrelatedness" of these problems, their causes, implications, and results, must be constantly kept in mind. Any one of them if isolated or treated alone could be either underrated or over-emphasized. With this consideration in view, then, the remaining questions raised in the Introduction will be answered, and some implications stated. Different interpretations may be advanced by other readers, but in the light of this investigation and the analyses of the data, these conclusions and implications seem justified.

b. Can any of the sources or causes of the difficulties be traced to the pre-service education programs of the normals or colleges?

If careful selection of candidates has been practiced, then some of the sources of trouble for these teachers seems to be traceable to these institutions. If selection has not been in effect, then the colleges are definitely a source of the difficulty, for they have recommended these people as being able and ready to teach.

The evidence shows that none of these institutions do, in general, any better job for their purpose than any other, and that their graduates all do better in knowing their subject matter than they do in controlling a school situation. This implies that (1) the colleges think subject matter more important than ability to control; or that (2) they are

[14] Even under more difficult conditions, these elementary teachers were not ultimately denied certificates in much greater proportion than were those teaching in high schools. For 15.8% of elementary, and 12.5% of high school teachers, were found in the "certificate denied" classification.

ignoring the possibilities of teaching, and giving practice in, this aspect of teaching; or that (3) "control" is something that can not be acquired by some people without actual practice in the field.

The theory that "control" is usually the most difficult aspect of teaching for beginners is borne out by the evidence here presented; and the implication seems clear that teacher-education programs should consider this fact more carefully and devise ways and means of overcoming this problem. This study emphasizes another point which is acknowledged in theory by educators: namely, that there are no real differences between ability to control and the other aspects of teaching (except knowledge of subject matter). This implies that it must be more clearly recognized that the "teaching act" is made up of a pattern of various factors which are all a part of the entire personality, attitudes, knowledges, and skills of the teacher, and that anything which effects a teacher's personality will effect her "teaching act" and, in turn, be effected by it. Making the implication more specific, this means that the training institutions should do more to help the teacher develop *as a person* in order to help him become a good teacher. Until this is more generally done, the colleges which train teachers must bear some of the blame for their failures.

c. Are the "rating criteria" and the raters themselves a source or cause of the failures of these teachers? For the cases in this study, the raters were all supervisory officials; therefore this really means: are the people in charge of the in-service program of education any source or cause of the difficulties?

The results of this investigation show that there is very little, if any, consistency in the judgment of the raters on the five criteria, and implies that there is some lack of ability of these people to rate. This deficiency of judgment may be due to a lack of understanding of the meanings or concepts of the terms employed; but if this is so, some definite effort should be made to clarify and "objectify" the meanings.

An examination of the records in the state department shows that there are many disagreements among officials regarding the recommendation of candidates for certification. Evidence is also there, in letters from teachers, showing that many of these people did not know that they were being rated for purposes of certification, or what criteria were being used for such evaluation. There are also statements which say, in effect, "If I had known I was failing in this respect, I would have tried to improve." If even a few of these inconsistencies, disagreements, letters, and statements are true, it would indicate that the supervisory officials are not giving the help they should to beginners, and are not considering carefully enough the merits and demerits of this rating system.[15] In these respects, then, the in-service education program seems of necessity to be included as a source of difficulty for these teachers.

3. Could any of the sources of the difficulties be in the communities to which these candidates go?

This has already been discussed in part, and the reply seems to be "yes" for these candidates, especially for the elementary teachers and particularly for the ones who had rural training and went to the open country to teach. The majority of these teachers went to communities of less than 2,500 population, and 158 of them (or 16%) went to 71 (or 14%) of the communities for part of their extended probationary period. Of the 14 most frequently cited communities, 9 of them had less than 1,000 population. The evidence seems to show that some of these small towns are difficult (if not "dangerous") places for beginners, and bears out the contentions of those writers who have said that ". . . over 50% of teacher failures find their direct cause in social maladjustment in the community,"[16] and ". . . community controls and expectations are common only in smaller and more isolated areas . . ."[17]

No fixed patterns or rules can, nevertheless, be described for all small towns or open country situations. Communities usually define their own school concepts, and these may deviate greatly from place to place, and from year to year. However, it seems clear that the implications should be restated that (1) communities should be selected for beginning teachers, as well as teachers for the community; and that (2) communities need to be educated to what their responsibilities are toward teachers. Underlining this whole concept is the theory that the people in a com-

[15] See Scully, Emily M., *Personal Administration in Public Education*, Ph.D. Thesis, University of Wisconsin, 1945.

[16] Nelson Bossing, *Progressive Methods of Teaching in Secondary Schools* (Boston: Houghton Mifflin, 1942), p. 67.

[17] L. A. Cook, *Community Backgrounds of Education* (New York: McGraw-Hill Book Co., 1938), p. 306.

munity want the best possible education—in its broader sense—for their children; and that to attain this, teachers must feel secure and adequate, and have a real and respected place in the community.

e. Are all of these teachers really failures?

This question has been dealt with in the preceding chapter. Only 14% of these 265 cases are really "legal" failures up to the present time. It remains to be seen how many in the "pending" classification will become so. Failure to be certified at some time or other, then, does not always designate one as a failure in the profession of teaching. The implication from this statement may be that considered judgment *might* have eliminated some of these "temporary" failures, and found the potential or real failures before too many children suffered at their hands.

f. What are the apparent results of these difficulties with certification? Are there any intangible ones which seem important and pertinent to this study?

Some of the answers to these questions have been implied throughout the discussion, but a brief summary will be made at this point of some of the more important ones.

1. The people who failed, even temporarily, to get a certificate at the usual time and under present practices, are faced with the prospect of worry, defeat, humiliation, and loss of status (in its personal and sociological sense). Consequently they are, in most cases, unable to do their best work, and not only they, themselves, suffer, but also the children with whom they are working. This implies that if teachers are inadequate for their jobs, they should be eliminated early; and if all they need is more time, experience, maturity, aid and so on, that deferment of certification be handled in such a way as to cause little, if any, humiliation and worry.

2. The teachers who failed permanently have been allowed to waste part of the school years of their pupils, to waste from four to six years of their own lives, and in some cases to become warped individuals for life. Aside from these factors—which some might call "sentimental" ones—there is the cold, hard fact of dollars and cents. What is the "money cost" of these failures for the people of the state? What is the "time" cost?

It is granted that a few such cases will occur, no matter how carefully the selection, training and supervision are handled. But this study has shown that many of these cases need not have occurred if a clarification of purposes, practices, and concepts took place between certification and supervisory officials, training institutions, placement services, and communities.

II. What are the effects of present methods of certification in Wisconsin upon teacher education and the teachers themselves?

Questions of this sort confirm the statement made several times previously: that all of these factors are so interrelated that they cannot be considered in any isolated manner. Yet the practice actually has been to separate in large measure the administrative aspects of certification from its educational implications. For, except in rare cases, the certification denials are handled automatically by referring to reference blanks and sending out form letters such as the one quoted in Chapter 1.

The effects of the present methods[18] of certification on teacher education have been negative rather than positive. That is, instead of recognizing and working toward better and more active participation with the teacher-education program, current practices in certification seem to be ignoring the possibilities of making the most of the relationships that do exist. These relationships were pointed out in the Introduction.

A definite "tying in" of the entire processes of pre-service education, licensing, recommending by in-service training personnel, further licensing, and granting or refusing a certificate, needs to be done. These are not isolated procedures for protecting children from poor teachers. That is one of the purposes of certification, to be sure, but that is only the negative aspect of the whole situation. A more positive approach is implied here: that of using these agencies for better service to beginning teachers by making the certification program a part of the pre- and in-service program for the improvement of teaching, and especially of assistance to beginning teachers. Other results of the present practices of this phase of certification have been stated and implied many times throughout this paper.

III. What are the implications of these findings for the improvement of certification and teacher education in Wisconsin?

[18] Note that the terms methods and practices are used, not regulations, standards and so on. The latter have had splendid results in lengthening the time of preparation, centralizing authority and responsibility, and beginning a new system of certification that is most promising.

This question brings us directly to the concluding recommendations.

Considering the results of this investigation and the conclusions and implications that have been drawn from them, the following recommendations for the improvement of certification and education of teachers are made:

1. That the State Superintendent organize an advisory committee to continue the study of the problems that have been raised here, and others of similar nature which pertain to the adjustment of beginning teachers to their schools and communities; and that this committee be composed of representatives of teacher-training institutions, supervisory officials, and certification authorities.

2. That this committee approach the study with (a) the concept of the beginning teacher *as a person,* who has to make multiple adjustments to a new community, new supervisory officials, new fellow workers and pupils, as well as to a new job; and with (b) the assumption that it is a responsibility of the supervisory and certification officials (including the pre-service personnel) to give more help to these teachers with their problems.

3. That qualifying for a certificate be recognized as a personal, social, and professional problem and not alone as a legal one; and that services to beginning teachers be coordinated by improving the relationships between college faculty members, certification authorities and supervisory officials.

4. That the criteria which are to be used for evaluating the work of beginning teachers be adopted by joint action of leaders in the pre- and in-service education programs and the legal certifying agency; and after they are decided upon, that the State Department should:

a. Send to every beginning teacher, every teacher-training institution, and every supervisory official a letter, giving in detail the criteria to be used for evaluation purposes. Clarify and describe the meaning of the criteria; and explain how and why they will be used.

b. Set up a policy and outline practices for a thorough review of all cases which are recommended for denial of certification, *before* writing a letter to the candidate.

c. Refuse to disqualify a candidate on the recommendation of only one, (and possibly, two) persons. Perhaps refuse also to grant a life certificate on the same basis.

d. Send to all teachers, at the close of the first probationary year, a letter, describing the record of that teacher to date, and pointing out matters that need improvement. Offer the help of state supervisors, if such help is requested.

e. Consider carefully, with other educators, the feasibility of lengthening the probationary period, especially for those teachers who have had less than four years of training. (While some might say that the five-year certificate is, in essence, permitting an extension of probationary time, the same attendant ills follow its denial as follow the refusal to grant a life certificate.)

f. Ask for references on candidates from the faculties of teacher-education institutions and supervising teachers as well as from superintendents, principals, and state supervisors. This implies visitation by faculty members.

5. That those responsible for "placing" candidates, study the communities to which young teachers go, and prepare both communities and teachers for meeting community-social and personal-social problems during the adjustment period.

6. That the entire teacher-education program clarify its concepts of training, supervision, and qualification for certification, and integrate the actual and potential relationships of these functions into vital, purposive, and human services for beginning teachers.

FACTORS AFFECTING REGULARITY OF THE FLOW OF WORDS DURING WRITTEN COMPOSITION

JOHN ANDREW VAN BRUGGEN
Grand Rapids, Michigan

SECTION I

INTRODUCTION

PROBLEM OF THIS INVESTIGATION

It is a matter of common observation that the flow of words during written composition is not regular and continuous. The periods of writing are irregular as well as the periods of reflection. Ideas seem to come at varying rates. The aim of this investigation is to obtain a quantitative measure of the rate of flow of words during written expression and to determine what factors affect it. This general aim will be achieved by comparing the rate of flow of words during the writing of each of two compositions by junior high school students with various test scores, ratings, and personal data.

The problem will be considered in four divisions. In the first a study will be made of the factors within the compositions that may affect the rate of the flow of words. Such aspects of the problem as writing on familiar and unfamiliar material, linguistic constructions, length, and general excellence of compositions will be considered.

The second division will treat the relation of academic achievement in the school subjects involving language skills to the rate of flow of words. Consideration will be given to the pupils' achievement in reading, spelling, and various phases of language.

In the third division several personal and environmental factors will be discussed. An attempt will be made to find the effect on flow of words of chronological age, grade in school, intelligence, perseverance, personality traits, handedness, eye dominance, sex, and such environmental factors as socio-economic status and position among other children in the family.

The fourth division will include analyses and comparisons of individual records. A detailed analysis will be made of the records of the nine compositions with the highest rating for quality and of the nine compositions with the lowest rating. The results will be compared to determine differences in flow of words between good and poor compositions. The records of the five compositions with the lowest indices of word flow and of the five compositions with highest indices of word flow will be analyzed and compared to find any relationship that may exist between the ability to write in thought units and the rate of flow of words.

The problem of this investigation is (1) to study the rate of flow of words during the writing of compositions to determine how it is affected by various compositional, academic, personal, and environmental factors; and (2) to analyze the records of flow of words for the individual pupils to discover how the composing structure—that is, the number, length, and location of pauses between words—differs in compositions of superior and inferior quality and in compositions written with rapid and slow flow of words.

SECTION II

METHOD OF INVESTIGATION

SOURCE OF DATA

The subjects.—The compositions of eighty-four pupils in the junior high grades of the Oakdale Christian School in Grand Rapids, Michigan, were used as the basis for this investigation. To keep constant the factors pertaining to school training, such as English curriculum and type of teaching, only those pupils were selected who had attended this school without interruption through all the grades, beginning with the first grade.

The subjects were distributed according to grades as follows: eight girls and thirteen boys from the seventh grade, twenty-three girls and nineteen boys from the eighth grade, and eleven girls and ten boys from the ninth grade. Originally another ninth grade girl was included, but the record of her writing proved faulty and had to be discarded. This left forty-two girls and forty-two boys.

The chronological ages of the subjects ranged from 146 months to 182 months with a median of 163 and a standard deviation of 9.18.

The compositions.—Each pupil who participated in this investigation was asked to write at three different sittings during May and June of 1943. Approximately ten days elapsed between writing periods. The writing was done in the music studio of the school while the subject was alone and free from the distractions of others.

During the first writing period the pupils wrote from memory the first seven sentences of Lincoln's Address at Gettysburg. This selection had been assigned for memorization to the junior high literature classes shortly before the time of this writing, and, consequently, it allayed any suspicions or fears of those who were asked to participate. Inasmuch as this first session was intended merely to give a record that could be used for determining speed of actual writing, it provided a splendid opportunity for practice.

Only two pupils were called to the place of writing at the same time. While one was writing in the studio, the other was in an adjoining room where he was asked by an examiner to recite the Address at Gettysburg. The few who could not recite fluently were instructed to spend some more time on the selection, and they were called back a few days later.

During the second period of writing the subjects were asked to reproduce the story of Androcles and the Lion. In preparation for this writing each pupil, upon entering the waiting room, was given a typewritten copy of the story and was instructed to read it carefully for the purpose of reproducing it. After the reading his comprehension was tested by an examiner. He was asked to answer orally fifteen questions dealing with the content of the story.

It was assumed that when pupils reproduced a story immediately after reading it, they were writing familiar material. The second selection of writing, which was the first composition of the subjects, was considered, therefore, as a sample of writing on familiar material.

For the third writing period the subjects were instructed to write a composition on one of six topics. A typewritten sheet with the following instructions was handed to the subjects as they entered the waiting room:

Here is a list of six topics. You may select any one of them for a composition. You may change the wording in some of them if you wish.

You will be given five minutes to decide on your topic and to plan your writing.

1. My Most Embarrassing Experience.
2. An Incident of My Vacation.
3. A Dream.
4. The Best Story I Ever Read.
5. A Most Exciting Experience.
6. The First Time I Went Fishing (Hunting, Bowling, etc.).

This third writing selection, the second composition, was regarded as a sample of writing on unfamiliar material.

Nature of Data

Records of the regularity of the flow of words.—Records of the regularity of the flow of words during the writing of the three selections were obtained by means of a motor driven, time recording kymograph. This apparatus is so constructed that a motor driven roller draws a strip of paper two inches wide over the flat surface or table of the machine. A punch set over the top of the machine is connected with a synchronous motor and a magnetic coil. A disc on the synchronous motor has fifty evenly spaced wires drawn over its outer edge. These are fastened to a copper ridge near the center of the disc. A copper spring connected to the magnetic coil scrapes along the ridge as the disc turns and connects the wire with the coil.

Two small rollers, which are connected with the coil, extend from the machine to the outer edge of the disc and touch the wires as the disc turns. The rollers are spaced so that when one touches a wire, the other is half way between wires. When the roller touches the wire, contact is made with the coil, and the coil draws the punch down to make a depression on the paper. Since the synchronous motor makes one revolution per minute and there are fifty strands of wire around the outer edge of the disc and each strand is touched twice during every turn, contact with the magnetic coil is produced one hundred times per minute. This means that a depression is made on the paper every six-tenths of a second as the paper is drawn through the machine.

A pencil set in a swinging bracket, that extends from the table of the machine over the paper, makes a line on the paper as it is drawn across the table. The bracket is shifted by another magnetic coil that is connected with a copper stylus and plate. When the stylus touches the plate and makes contact, the bracket shifts so as to make an indentation in the line that is being drawn on the paper.

The machine was located in a room across the hall from the studio in which the subjects did their writing. The stylus and contact plate were in the studio and were connected with the machine by wire. This arrangement eliminated the possibility of the subject's hearing the noise of the machine.

In the studio the pupil who did the writing was watched by the examiner, who was seated at a small table with the stylus and contact plate. Between her table and the desk of the pupil was an opaque portable screen in which was mounted a piece of one-way glass about a foot square. Through this glass the examiner could watch the writer without being observed. The subject was conscious of the examiner's presence but could not see her and did not know what she was doing. The examiner got the subjects started at the writing assignment by providing pencil and paper and asking them to let her know when they started writing and when they were finished. She then went to her place back of the screen.

When the pupil placed his pencil on the paper to begin writing a word, the examiner placed the stylus on the contact plate thereby shifting the pencil on the apparatus and indenting the pencil line that was being drawn on the record paper as it passed over the table of the machine. When the writer lifted his pencil at the end of a word, the examiner raised the stylus to break the contact and to shift the pencil on the apparatus back to its original position. If the writer stopped to cross out a word, the crossing out was recorded by the examiner as a word. The pupils had been instructed beforehand to cross out words rather than to erase them. When the pupil came to the end of a line, the examiner touched the contact plate with the stylus three times in rapid succession. This left a mark on the paper that facilitated interpretation of the record.

After the subject had finished writing, his record paper was removed from the machine and his name written on it. The record paper showed a series of depressions down the middle to indicate time in six-tenths of seconds and a broken pencil line along the edge to indicate periods of writing and pauses between words.

The first step toward using the records involved the transfer of the words of the composition to their appropriate indented spaces on the pencil line of the record. This was done quite readily by comparing the paper on which the subject did his writing with his record. The three marks designating the end of a line made it possible to transfer the words of one line at a time.

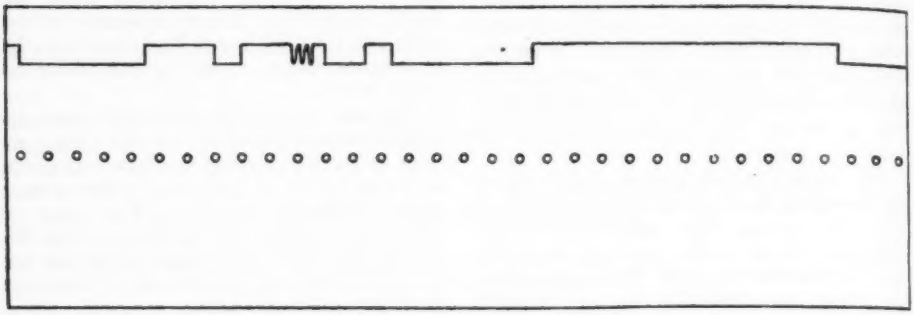
The next step was to determine the length of time used for each word and the pause between words. At each point in the record where the pencil line was broken, another line was drawn across the paper at right angles to the record line. By counting the number of round depressions between these lines and multiplying by six-tenths, the time in seconds was obtained for each word and each pause. The three stages in the interpretation of the record are shown in Plate I, which is a reproduction of part of the third record of Subject One.

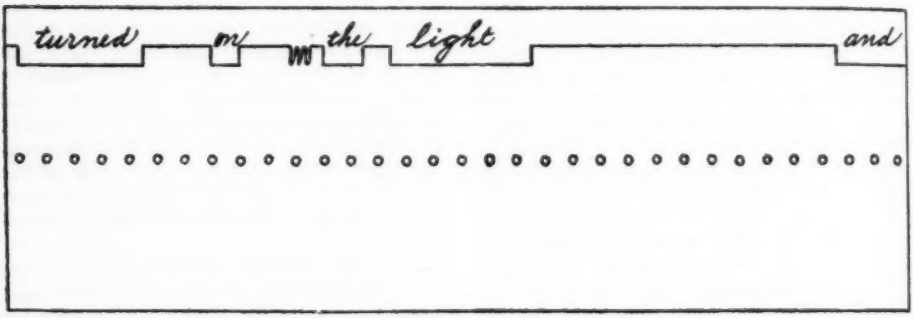
After this had been completed, the average time required for writing each word was found by adding the number of seconds for each word and dividing the total by the number of words.

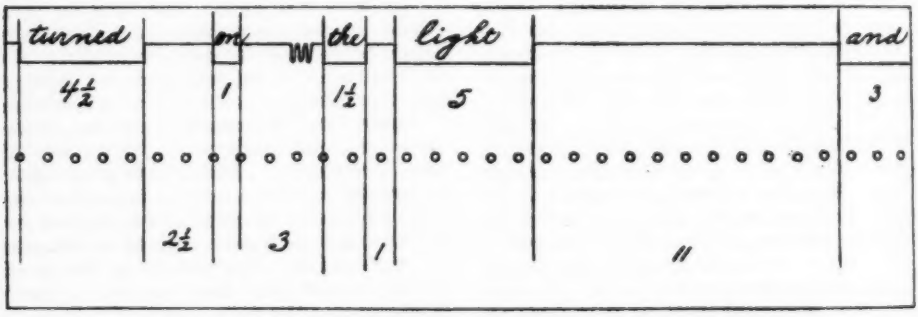
Finally, a distribution was made for each subject showing the number and length of the pauses in each of the two compositions. No distribution of pauses was made for the first record of writing since it was a reproduction of a memorized selection and a record of the pauses was not pertinent to the purpose of this investigation. The distribution sheet showed the number of words in each of the three selections, the average time required for the writing of each word, and the distribution of pauses for the two compositions. A copy of the distribution sheet that was used is shown in Figure 1.

Standardized test scores.—During December of each year standardized achievement tests are administered to all the pupils of the Oakdale Christian School as part of the school's regular testing program. Form E of the Stanford Achievement Test: Advanced

PLATE I







A REPRODUCTION OF PART OF A RECORD ILLUSTRATING THE THREE STEPS REQUIRED FOR INTERPRETATION.

DISTRIBUTION SHEET

Name_____ Grade_____

First Record		Second Record		Third Record	
Number of words:_____		Number of words_____		Number of words_____	
Seconds per word:_____		Seconds per word_____		Seconds per word_____	
Pauses:		Pauses:		Pauses:	
.3		.3		.3	
.6		.6		.6	
.9		.9		.9	
1.2		1.2		1.2	
1.5		1.5		1.5	
1.8		1.8		1.8	
2.1		2.1		2.1	
2.4		2.4		2.4	
2.7		2.7		2.7	
3.0	13.8	3.0	13.8	3.0	13.8
3.3	14.1	3.3	14.1	3.3	14.1
3.6	14.4	3.6	14.4	3.6	14.4
3.9	14.7	3.9	14.7	3.9	14.7
4.2	15.0	4.2	15.0	4.2	15.0
4.5	15.3	4.5	15.3	4.5	15.3
4.8	15.6	4.8	15.6	4.8	15.6
5.1	15.9	5.1	15.9	5.1	15.9
5.4	16.2	5.4	16.2	5.4	16.2
5.7	16.5	5.7	16.5	5.7	16.5
6.0	16.8	6.0	16.8	6.0	16.8
6.3	17.1	6.3	17.1	6.3	17.1
6.6	17.4	6.6	17.4	6.6	17.4
6.9	17.7	6.9	17.7	6.9	17.7
7.2	18.0	7.2	18.0	7.2	18.0
7.5	18.3	7.5	18.3	7.5	18.3
7.8	18.6	7.8	18.6	7.8	18.6
8.1	18.9	8.1	18.9	8.1	18.9
8.4	19.2	8.4	19.2	8.4	19.2
8.7	19.5	8.7	19.5	8.7	19.5
9.0	19.8	9.0	19.8	9.0	19.8
9.3	20.1	9.3	20.1	9.3	20.1
9.6	20.4	9.6	20.4	9.6	20.4
9.9	20.7	9.9	20.7	9.9	20.7
10.2	21.0	10.2	21.0	10.2	21.0
10.5	21.3	10.5	21.3	10.5	21.3
10.8	21.6	10.8	21.6	10.8	21.6
11.1	21.9	11.1	21.9	11.1	21.9
11.4	22.2	11.4	22.2	11.4	22.2
11.7	22.5	11.7	22.5	11.7	22.5
12.0		12.0		12.0	
12.3		12.3		12.3	
12.6		12.6		12.6	
12.9		12.9		12.9	
13.2		13.2		13.2	
13.5		13.5		13.5	

Fig. 1. Sheet Used to Show Distribution of Pauses.

Battery, Partial, was administered to the junior high pupils in December, 1942, approximately five months prior to the time of the writing for this investigation. The scores of the Reading: Paragraph Meaning, Reading: Word Meaning, Language Usage, and Spelling sections of the test were used in this study.

The following tests were administered during February and March of 1943 to the classes participating in this investigation.

The Terman McNemar Test of Mental Ability: Form C.

Co-operative English Test, Test B1: Effectiveness of Expression, Lower Level, Form Q.

Co-operative English Test, Test A: Mechanics of Expression, Form Q.

Aspects of Personality Test, for Grades 4 to 9 Inclusive, by Pintner, Loftus, Forlano, and Alster.

Carroll Prose Appreciation Test for Junior High School.

All of these tests were administered and scored by the writer, and the scores were used in this investigation.

Personal data.—Other information pertaining to the subjects was obtained from the school records. This included the occupation of the father, number of children in the family, number of children older and younger than the subject, chronological age, handedness, and marks in English.

Eye dominance was tested in two ways. A paper funnel with the wide end fitted to the contour of the face was used in the first test. A ball was placed on a shelf about twenty feet away from the subject. He was then asked to look through the funnel with both eyes and locate the ball. After he had located it, a card was inserted in a slot in the funnel so that the left eye was covered. If the subject could still see the ball, he was regarded as having right eye dominance. If he could not see it with his left eye covered, but could see it with his right eye covered, he was regarded as having left eye dominance.

The second test was a card test. A large card about twelve inches square with a small hole about an eighth of an inch in diameter in the middle was given to the subject, and he was asked to look through this hole at the ball on the shelf. Placing the card in front of and looking with the right eye was regarded as an indication of right eye dominance, and looking with the left eye was regarded as an indication of left eye dominance.

Twelve of the subjects showed a different eye dominance on one test than on the other. They were requested to take both tests again, and in each case the results of the two tests agreed. This lack of agreement in the first testings may have been due to a failure on the part of the subjects to focus correctly when using the funnel.

Method of Tabulating Data

Finding a measure of the regularity of the flow of words.—After a distribution had been made of the lengths of the pauses between words in each composition, it was necessary to find some measure of the variability of the pauses. This measure could then be used as the measure of the regularity of the flow of words.

The commonly accepted measure of variability is the standard deviation, and it was computed for each of the compositions. They ranged from .28 to 4.73 with a median of 1.26 for the first set of compositions, and from .29 to 8.76 with a median of 1.31 for the second set of compositions. The coefficient of correlation between the standard deviations of the first and second sets of compositions was .48 with a probable error of .06.

Since the standard deviation represents the square root of the squared deviations from the median, its use might emphasize unduly the extremely long pauses in some of the compositions. It was necessary, therefore, to cast about for some measure that would take into account the long pauses without weighting them disproportionately.

The average of all the pauses longer than the median pause, or the average of the higher half of the pauses, was computed for each composition. These averages ranged from .67 to 4.72 with a median of 1.80 for the first group of compositions, and from .80 to 4.33 with a median of 1.89 for the second group. The Pearson product-moment coefficient of correlation was computed for the averages of the first and second sets of compositions and was found to be .77 with a probable error of .03. This is .29 higher than the coefficient of correlation of the standard deviations of the two sets of compositions. If regularity of flow of words is a trait that should be fairly constant in all writing, then the higher coefficient

of correlation would indicate that the average of the higher half of the pauses is a better measure of the flow of words than the standard deviation.

The coefficient of correlation between the standard deviations and the averages of the higher half of the pauses was .91 with a probable error of .01 for the first group of compositions and .92 with a probable error of .01 for the second group. These high coefficients of correlation as well as the high correlation between the sets of compositions seem to justify the use of the average of the higher half of the pauses as a valid measure of regularity of the flow of words. It will be used, therefore, in further comparisons and will be referred to as the index of flow of words. A high index indicates a slow flow of words, and a low index indicates a rapid flow of words.

Comparing data.—The Pearson product-moment coefficient of correlation was used to determine relationship between the indices of flow of words and the other data compiled for this investigation.

Additional light on the relationships was obtained by arranging the scores of the eighty-four subjects in each of the areas investigated in order of rank and dividing them into four equal groups. The mean index of flow of words was then found for each group, or quarter, of the scores. This method proved to be enlightening especially in the areas in which the distribution of scores tended to be curvilinear.

SECTION III

FACTORS WITHIN THE COMPOSITIONS AND THEIR EFFECT ON THE FLOW OF WORDS

In a study of sentence structure in children's compositions Shepherd called attention to content as an important factor in determining sentence structure.[1] Inasmuch as content might also affect the regularity of the flow of words, the compositions submitted by the subjects were carefully analyzed to ferret out any factors that might affect word flow.

Familiar and Unfamiliar Material

Writing speed.—In an effort to discover any significant changes in the writing speed in the three selections of each subject, a study was made of the average time required per word in the actual writing of each selection.

In the first sample of writing, the portion of Lincoln's Address at Gettysburg, the average time in seconds per word required for actual writing ranged from 1.21 to 2.66 with a median of 1.81 and a standard deviation of .32. In the second selection, the reproduction of a story, the time per word ranged from 1.02 to 2.62 seconds with a median of 1.59 and a standard deviation of .29; and in the writing of compositions on selected topics the time per word ranged from 1.24 to 2.63 seconds with a median of 1.62 and a standard deviation of .24.

Seventy-three of the subjects took more time for the actual writing in the first selection than in the second, and sixty-eight took more time per word in writing the first selection than in writing the third. The long and unfamiliar words in the Address at Gettysburg may account for part of this slow writing rate, but there is enough difference to indicate that in writing a selection from memory the need for recalling exact wording retards writing speed.

Since writing to reproduce a story is more likely to involve the use of unfamiliar words than writing on a selected topic, the writing rate of the first composition might be expected to be slower than that of the second. Yet forty-eight, or fifty-seven per cent, of the subjects took more time per word in the second composition than in the first. The median time per word was also longer for the second set of compositions than it was for the first.

There is evidence here to suggest that when pupils write on unfamiliar material and must compose as they go along, as they did in the second composition, they have a tendency to slow down in their actual writing rate.

Regularity of the flow of words.—A comparison was also made of the indices of the flow of words in the two compositions. They ranged from .67 to 4.72 with a median of 1.80 and a standard deviation of .73 for the first group of compositions and from .80 to 4.33 with a median of 1.89 and a standard deviation of .78 for the second group. The wide range in the first group of compositions is the result of a few extremely high indices, for there were only six indices above 2.80 in the first group while there were thirteen in the second.

[1] Edith Elizabeth Shepherd, *The Influence of Content and Maturity on Sentence Structure in Pupils' Written Discourse.* Unpublished Master's thesis, Department of Education, University of Chicago, 1933.

Fifty-one, or 61 per cent, of the pupils had a higher index of flow of words in the second composition than in the first; but the coefficient of correlation between the indices of the two sets of compositions was .77 with a probable error of .03.

The higher indices of flow of words for the second group of compositions as revealed by their range, median, and standard deviation as well as the fact that the majority of the subjects had a higher index in the second composition than in the first imply a retardation in flow of words when pupils write original compositions that must be composed as they write.

Writing speed and the flow of words.—The average time per word for actual writing and the index of flow of words in each of the two sets of compositions were compared to bring to light any relationship that might exist between them.

The coefficient of correlation between writing speed and the indices of flow of words was found to be .34 with a probable error of .07 for the first group of compositions and .29 with a probable error of .07 for the second group. These correlations are sufficiently high to indicate a positive relationship between fast writing and rapid flow of words.

The compositions were arranged according to writing time from the highest number of seconds per word to the lowest, and then divided into four equal groups. The mean index of flow of words was then found for each group, or quarter, of the compositions. The results are given in Table I.

TABLE I

THE MEAN INDICES OF FLOW OF WORDS FOR EACH QUARTER OF COMPOSITIONS ARRANGED ACCORDING TO WRITING SPEED

Writing Time	Mean Index of Flow of Words	
	First Set of Compositions	Second Set of Compositions
Highest Quarter	2.22	2.58
Second Quarter	2.01	1.84
Third Quarter	1.81	1.81
Lowest Quarter	1.67	2.00

In the first set of compositions the mean indices decrease regularly with a corresponding decrease in writing time per word. The same trend is apparent in the second set of compositions, although it is not as regular and consistent. The trend is clear enough to indicate, however, that as the flow of words increases the speed of writing also increases. This is especially true in compositions involving writing on familiar materal.

The results point to the probability that some pupils have a slow reaction time that affects their rate of writing as well as their flow of ideas.

Other compositional factors that were considered in their relation to the flow of words were the prevalence of certain linguistic constructions, the length, and the quality of the compositions.

SUMMARY AND CONCLUSIONS

In an investigation such as this all the factors that may be presumed to be related to flow of words have to be considered. Obviously, not all of them will prove to be related. In some cases the relationship may be so little as to result in practically zero correlation.

The comparisons between flow of words and some of the factors studied may indicate a relationship that is merely indirect. These factors may be related because of their high correlation with other factors that are related to flow of words.

The conclusions at the end of each section will deal primarily with the more important relationships revealed by the comparisons made in that section. A discussion of the significance of these relationships will be reserved for the last section.

Inasmuch as the number of subjects whose compositions were used in this study is comparatively small, the conclusions should be interpreted as suggestive of trends rather than as conclusive evidence of relationships.

In this section the three samples of writing submitted by each subject were carefully studied and compared to bring to light any factors within the compositions that might affect the rate of the flow of words during writing. A brief summary of the results is presented in the statements that follow:

1. A comparison was made of the time in seconds per word for actual writing in each of the three selections of each subject. The rate of writing in the first selection, writing from memory, was slower than that of either of the two compositions. Of the two sets of compositions the second, which involved writing on selected topics, showed a slower writing rate than the first, a reproduction of a story. The difference between the rates of

the two compositions, however, was not so great as the difference between the rate of the first selection and either of the compositions. These comparisons point to the conclusion that the necessity of recalling exact wording and the need of organizing material during the process of writing affect writing speed adversely.

2. A comparison of the rates of flow of words in the two sets of compositions revealed that on the whole the pupils hesitated more while writing on selected topics than they did while writing to reproduce a story with which they were familiar.

3. The writing speed and rate of flow of words were compared for each group of compositions. The comparisons showed that rapid flow of words and rapid writing went together with a tendency to decrease and increase concomitantly. This tendency was not as clear in the second group of compositions as it was in the first.

4. The second group of compositions was carefully analyzed for the presence of linguistic constructions that reveal dominant traits of the writers. The four types of constructions considered are: (1) the Adjective-Verb Quotient, denoting emotional stability; (2) the Index of Subordination, signifying linguistic maturity; (3) the Personal Pronoun Index; and (4) the First Person Pronoun Index, both showing the extent of social development.

The comparisons showed no relationship between emotional stability, as measured by the Adjective-Verb Quotient, and rate of flow of words. The comparisons did show an apparent tendency for those pupils who were on the higher and lower ends of the scale of linguistic maturity or the scale of social development to write with a slower flow of words than those near the middle of the scale. Those on the higher end, however, had a better flow of words than those on the lower end.

5. Composition length, as determined by the number of words, was compared with the rate of flow of words for each of the two sets of compositions. In the first set those who wrote compositions of medium length had the best rate of flow of words, but in the second set those who wrote the longest compositions had the best rate. Those who wrote short compositions had a comparatively slow rate of flow of words. This was especially true in the second group of compositions.

6. The compositions were rated by three junior high English teachers. The mean coefficients of correlation of their ratings were .70 for the first set and .58 for the second set. The distribution of the ratings of the teachers shows that while the marks of the second set of compositions are spread out more over the whole range of quality, than those of the first set, the quality of the compositions in the second set, on the whole, is better than that of the compositions in the first group.

The sum of the point values of the three ratings was taken as the measure of quality and compared with their corresponding indices of flow of words. The comparisons for the first set of compositions showed no relationship between quality and the rate of flow of words. The comparisons for the second set showed a tendency for the pupils who write compositions of very inferior quality to hesitate a great deal as they write.

7. The comparisons reveal that a relatively slow flow of words may be expected when pupils write on unfamiliar material, when they are at the higher and lower ends of the scales of linguistic maturity and social development, especially at the lower end, and when they write compositions that are extremely short or of very inferior quality.

SECTION IV

ACADEMIC FACTORS AND THEIR EFFECT ON THE FLOW OF WORDS

Introductory Statement

Linguistic ability is commonly considered to be in part a result of academic training. It is frequently assumed that training in those school subjects that involve language skills will affect a pupil's ability to express his thoughts in writing.

The validity of this assumption may be tested by a study of the relationships between ability in these various subjects and facility of written expression.

The purpose of this chapter is to report on the phases of such a study. Ability in the fields of reading, spelling, and language were compared with facility of expression as measured by the rate of flow of words during composition writing.

Summary and Conclusions

In an effort to discover the relationship between ability in academic subjects involving language skills and the rate of flow of words during composition writing, comparisons were made between scores designating ability in those subjects and the indices of flow of words for the two sets of compositions. The scores included English marks for eight semesters as well as scores on standardized tests in reading, vocabulary, spelling, language usage, mechanics of expression, effectiveness of expression, and prose appreciation.

The first step toward finding the relationship between the scores in those subjects and the flow of words was to compute the coefficient of correlation between the indices of flow of words for each set of compositions and the scores on the standardized tests. A summary of the coefficients of correlation with their probable errors for each of the subjects is given in Table II.

Inasmuch as the correlations revealed little or no relationship between the test scores and the flow of words, other comparisons were made. The scores were arranged in order of rank from the highest to the lowest, and divided equally into four groups. The mean index of flow of words was then found for each of these groups or quarters.

The comparisons reveal:

1. That pupils who indicate by standardized test scores that they have low reading, vocabulary, and spelling ability also have a slow flow of words during writing.

2. That the best spellers show a tendency to hesitate between words when writing compositions, although their rate of word flow is not as slow as that of the poor spellers.

3. That a low level of ability in language, as indicated by school marks, appears to be accompanied by a rapid flow of words during composition writing. Pupils who are just slightly below average in language ability have a comparatively slow flow of words.

4. That pupils with high scores in language usage are apt to have a rather free and rapid flow of words during composition writing. This rate of word flow has a tendency to slow down as the pupils' scores become lower. Pupils who show a very limited knowledge of language usage, however, seem to pause very little between words as they write their compositions.

5. That the pupils with the best scores in mechanics and effectiveness of expression tend to have the most rapid flow of words. Those with low scores have a slow rate of word flow. The change from slow to rapid rate is not regular, however, and does not proceed concurrently with improvement in mechanics or effectiveness of expression.

6. That prose appreciation seems inversely related to rate of flow of words, for the pupils with the highest scores on the prose appreciation test had the slowest flow of words. Pupils with low scores also showed a slow flow of words, however; but their rate was not as slow as that of the pupils with the high scores.

7. That on the whole pupils with low scores in the school subjects involving language skills tend to have a slow rate of flow of words during composition writing.

8. That a high level of ability in school subjects that involve language skills does not guarantee a rapid flow of words. It may tend to retard word flow.

Table II

Coefficients of Correlation Between Scores on Achievement Tests and Indices of Flow of Words

Subject	First Compositions		Second Compositions	
	Coefficient of Correlation	Probable Error	Coefficient of Correlation	Probable Error
Reading:				
Paragraph Meaning	.03	.07	.03	.07
Word Meaning	.10	.07	.14	.07
Spelling	.04	.07	.13	.07
English marks	.00	.07	—.01	.07
Language usage	.00	.07	—.02	.07
Mechanics of expression	—.03	.07	.08	.07
Effectiveness of expression	.13	.07	.11	.07
Prose appreciation	—.09	.07	—.12	.07

SECTION V

PERSONAL AND ENVIRONMENTAL FACTORS AND THEIR EFFECT ON THE FLOW OF WORDS

Introductory Statement

Language is a part of every individual's development, but what language he uses depends on the group into which he is born.[2]

The language one uses is probably the basis upon which his general culture and personality are judged more often than any other single index. What one says and his way of saying it are interpreted almost universally as indicative of the kind of person he is.[3]

From these statements it is evident that language is considered as something personal and that a study of the language habits of any group must include a consideration of the individuals in that group and the environment to which they have been exposed.

In so far as rate of flow of words is a measure of linguistic behavior, it cannot be understood without a consideration of the personal and environmental factors involved. Several personal traits and measures of personal development were considered, therefore, in an effort to determine their effect on the rate of flow of words.

The personal factors included chronological age, grade in school, mental age, intelligence quotient, sex, perseverance, ascendance-submission, extroversion-introversion, emotional stability, handedness, and eye dominance. The environmental factors considered were socio-economic status, as determined by the occupation of the father, and place among other children in the family.

Summary and Conclusions

The coefficients of correlation were computed to determine the correlation between the various personal factors and the flow of words, but because they were low and because correlations do not take into account all the possibilities of relationship, a second group of comparisons was made.

The scores and measures of the various personal factors were arranged in order of rank and divided into four equal groups. The mean indices of flow of words were then computed for these groups and for other groups that designated personal traits. These comparisons revealed:

1. A definite relationship between chronological age and rate of flow of words. The rate of flow of words was found to increase with increasing chronological age.

2. That the rate of flow of words increases as the pupils pass from grade to grade. Ninth grade pupils had a better flow of words than eighth graders, and eighth grade pupils had a better flow of words than seventh graders. Since there is a high correlation between chronological age and school grade, these findings corroborate the conclusion that rate of flow of words is associated with chronological age.

3. That there is a relationship between mental age and flow of words. The pupils with the high mental ages showed a rapid flow of words and the pupils with the low mental ages showed a slow flow of words. The increase in rate of flow of words that accompanies rise in mental age is not as rapid as the increase that accompanies rise in chronological age during the junior high school years.

4. That pupils with very low and very high intelligence quotients seem to have the best flow of words.

5. That girls do not consistently have a better rate of flow of words than boys. The seventh grade girls had approximately the same rate of flow of words as the boys; but as the rate increased with the rise in school grade, the rates for the girls showed a more rapid increase than those for the boys.

6. That pupils who are rated high in perseverance are apt to have a rapid flow of words as well as those who are rated low. Those near the middle of the scale show a tendency to write with greater hesitation.

7. That the strong, dominating type of pupils as well as the meek and submissive type, as indicated by personality test scores, show a tendency to compose freely and without much hesitation. Those who fall between these two extremes may be expected to have a slower rate of flow of words during their composition writing.

8. That extrovertive pupils pause very little between words as they write their compositions, but as pupils tend toward introver-

[2] *Language in General Education*, A Report of the Committee on the Function of English in General Education for the Commission on Secondary School Curriculum (New York: D. Appleton–Century Company, 1940), p. 10.

[3] M. R. Trabue, *Teaching Language in the Elementary School*, Forty-third Yearbook of the National Society for the Study of Education, Part II (Chicago: National Society for the Study of Education, 1944), p. 1.

sion the number and length of pauses seem to increase.

9. That emotionally stable individuals write their compositions with a comparatively free and rapid flow of words. The rate of flow of words decreases as the individuals show increasing indications of emotional instability.

10. That pupils with right hand and right eye dominance and consistent hand eye dominance have a better rate of flow of words than others. Pupils with crossed hand eye dominance show a slow flow of words.

11. That the evidence regarding the relationship between socio-economic status and rate of flow of words is not very definite, but there is some indication that pupils who have socio-economic advantages write with a better flow of words.

12. That there is no relationship between rate of flow of words and position among children in the family. Only children, oldest children, youngest children, and others all showed approximately the same mean indices of flow of words.

SECTION VI

DETAILED ANALYSIS OF INDIVIDUAL RECORDS

Introductory Statement

Another phase of this investigation of flow of words during composition writing involves a study of the individual records of the subjects. An analysis of these records to determine the frequency and location of pauses may uncover some very significant facts.

It has already been pointed out in Section III that the pupils who wrote very inferior compositions showed great hesitation during their writing, as indicated by their high indices of word flow. It is important to know how the composing structure—that is, the number, length, and location of the pauses—differs from that of the good compositions. The records of each of the nine best compositions were therefore carefully analyzed and compared with the records of the nine poorest compositions to discover differences in length and location of pauses.

The question of whether rate of word flow is related to composing ability has not been definitely answered by the comparisons made so far in this investigation. If good composing, like good reading, requires the grouping of words into thought units, the number and location of pauses ought to reveal composing ability. The records of the five subjects with the highest indices of word flow were therefore examined for length and location of pauses and compared with the records of the five subjects with the lowest indices of word flow.

Analysis of the Records of Good and Poor Compositions

Records of the nine best compositions.—The records of the subjects with the nine best compositions were first analyzed for composing structure. The record of the best composition is reproduced for the purpose of illustration. The dashes between the words represent pauses, each dash representing a pause of three-tenths of a second. No dash is used when the pause is less than a half second.

Inasmuch as the indices of word flow vary with the different compositions, it was necessary to decide on some method of determining significant pause length for each composition. The following procedure was adopted:

A time length for significant pauses was arbitrarily selected. The pauses of that length and longer were then counted and classified as either those which separated units of thought or those which broke up units of thought. A pause length slightly longer was selected and the pauses were again counted and classified to determine any change in the proportion of the one kind of pause to the other. If there was no change, pauses of shorter duration were counted and classified. If the ratio remained the same, a smaller time interval was used and again the smaller pauses were classified. The time interval at which the ratio of pauses that break up thought units to pauses that separate thought units increased perceptibly was selected as the significant pause length for the composition. Only pauses of significant time length were considered in the comparisons.

The best composition was written by Subject 51. The record follows:

> Splash,..splash,..the dip...of the oars sent a thrill....to...my heart as the small row...boat pulled out away from the...dock. I was going fishing...for...the first..time.....
>
> Surrounded..by tackle and...hooks I cast my line into....the water..and sat waiting.....for the first..bite, which,...I...felt confident,.... would come in...a few minutes.....

Much time...elapsed and...still..not even a nibble..d r a g g e d...the..peacefully....bobbing ..float beneath....surface...."Remember,"..... ..my uncle...was saying,....."don't..stand..up ...when you catch a fish....You...might.... fall."..... .

His words..fell on deaf.....ears...My..bobber ..was acting.....very strangely....Then..suddenly...it went completely.....beneath...surface....Struggling.....to my...feet..and..waving ...my arms I..shouted,...."I've got a fish, I've ....got a fish."..... .

Then it..all..happened at....once. There..was a big..splash...and I felt..myself....sinking.... ..down, down,..into..the blue depths.....There ....was a moment of oblivion....and..then..I found...myself.....lying...p r o s t r a t e...on the bottom of the...boat.....My uncle was standing ....over me and shaking...his head....disgustedly..... .

We came that night...without....any fish..and ..I..felt that the...worst part of the enterprise .....was the lecturing...I got when...I got home. ...I resolved then..and there that..*I would*...if ...I ever went..fishing....a g a i n...I would.. never, never stand up..in the boat.

The record indicates that this composition was written with few extremely long pauses. The longest pauses occurred at the end of paragraphs, and the pauses of one and a half seconds and longer occurred consistently between groups of words that represent thought units.

Analysis of the records of the nine best compositions reveals great individual differences among the writers in composing fluency. Some write with few and short pauses while others hesitate frequently and for comparatively long periods of time.

The general trend of the pauses in these nine compositions, however, follows a definite pattern. The longer pauses generally occur between groups of words that represent thought units, and the shorter pauses occur within these groups of words. There is evidence to conclude that the good writers proceed by thought units, writing each unit as it comes to them.

The pattern is distorted when the writer indicates by the crossing out of words that he is not fully prepared for the writing or that he has not fully thought through his material. In such cases the writing is stopped and the words crossed out even before the entire thought unit has been written. From this it appears that the thinking is well in advance of the writing, and that the pupil, while writing a thought unit, is already framing the next.

Records of the nine poorest compositions.—The records of the nine compositions that were given the poorest ratings were also analyzed.

The record of the composition of Subject 75, given a rating of 7, follows:

One...New..Year's Day.....w h e n..I was.. just.....sitting.....around..the..h o u s e... and didn't....h a v e.....nothing...to..do..... .....I....planned to go..hunting..... ...out to...my Grandad's.....house..... .....When..I..left I took..... ..25.....shells....along for my.....410... shot gun.....I.....w a l k e d ..out *there*.....pa[illegible].....of the..way .....Then..... .when I..was..getting....tired...I....planned to .....start..hitch...hiking..... and I did.....I..got a..ride..out there.... .....When.....I got..there...I..... .....got...my..uncle's..dog out..and.. went..out.....into..the.....pasture..... .....Then I let the dog go.....a n d..find..some pheasants to...scare up,.....he did. And..... .....

.I wasted..5..shells on.....a pheasant..... .....b u t....I....got..his.. tail..feathers..... About ..an hour..later.....the..dog chased up .....a..rabbit.....I got..... him but..could..not...eat him because..... .....he..was full..... ....of...w o r m s.....Then the dog.... scared up.....the....same..pheasant...he..did ..the...first..time.....I got.....him.. .....right.....through....the....head .....Then....after...roaming..out..there.....about.....three..... hours.....and not a..thing..to shoot at... but a....few..s q u i r r e l s...I..planned to go.. home.....w h e n....I got up....to..my..grandad's....house..my..uncle..... .....w a s..going..to..town..... .....so...I rode..home..with...him..... .....

The..next.....n i g h t...we..had.. pheasant.....for supper.....Does.. it..ever taste.....g o o d....if..you...kill.... and. .....cook..it.....yourself.

This composition, with an index of word flow of 2.87, has thirty-five pauses longer than two and a half seconds. The original composition was written on both sides of a sheet of paper, and one of the pauses occurs at the place where the paper was turned over. Of the remaining thirty-four pauses, eleven break up units of thought and twenty-three separate thought units.

A comparison of the nine best and nine poorest compositions reveals that the better compositions are more definitely written in thought units. A summary of the analyses

made of the records of the eighteen compositions shows that in the nine best compositions only twenty-four pauses, or 17 per cent, occur between words of a thought unit and 115 pauses, or 83 per cent, come at the end of thought units. In the nine compositions with the lowest rating of quality forty-seven pauses, or 31 per cent, occur between words of a thought unit and 107 pauses, or 69 per cent, separate groups of words expressing units of thought. In short, in the compositions of poor quality the pauses that break up thought units occur approximately twice as often as they do in the compositions of better quality.

The comparisons also reveal that the pauses are longer in the poorer compositions. The average index of word flow is 1.6 for the best compositions and 2.3 for the nine poorest compositions. Of the pauses longer than two and a half seconds, 115 occur in the nine best and 165 in the nine poorest compositions.

The evidence is clear, therefore, that poor composition writing is accompanied by or the product of uncertainty of expression. This uncertainty is reflected in the breaking up of thought units and the use of long pauses.

Analysis of the Records of Compositions With High and Low Indices of Word Flow

Additional evidence on composing ability may be gained from a study of the records of the compositions with highest and lowest indices of word flow. An examination of these records will bring to light any relationship that may exist between composing speed and composing structure.

The records of the five compositions with lowest indices of word flow were analyzed and compared with the records of the five compositions with highest indices of word flow. The records and the graphs which illustrate the location of the pauses are given in the following pages.

The graphs illustrate the composing structure of the first fifty words in each of the compositions. The heavy horizontal line represents a continuous time-line in writing the composition, and the round black spots represent the words. Each of the ten heavy horizontal lines is divided into ten sections, so that each section represents 1 per cent of the total time of the pauses between the fifty words, for the spacing is expressed in per cent of total time rather than in actual time units. The numbers above the black spots correspond with the numbers above the words in the record of the composition.

Records of the compositions with the lowest indices of word flow.—The pupil who had the lowest index of word flow, .80, is Subject 79. The record of the composition follows:

1 2 3 4 5 6 7 8
Jean..and I were playing..on..her front..
9 10 11 12 13 14 15
porch when along..came Georgia..Lou...We
16 17 18 19 20 21 22 23 24
started...pushing her on the swing but...she soon
25 26 27 28 29 30 31 32 33
wanted..to..get..off....We..wouldn't let her but
34 35 36 37 38 39 40 41
...finally...she managed..to..run..off and ran
42 43 44 45 46 47 48 49 50
home yelling..back..that she'd get her daddy to
...shoot..us...Of course we were..scared stiff...
and..ran into..Jean's..house..to..hide..in..a..
closet..but her mother wouldn't..hear..of it...We
..ran to my....house but my..mother..was...
equally..unsympathetic....We....ran outdoors..
again thinking.....we wouldn't see the light
of....day again.....In our garage.....was our
car with..the doors...unlocked...We...went..in-
side..and..both lay on..the..b a c k..seat...for
Jean told..me not to lie....on the floor..because
the man....could feel that...We lay there..quietly
...scarcely..breathing.....All of a sudden we
heard....footsteps.....and our hearts...stopped
..beating.....They..walked..around in the garage
..and..finally..w e n t out....We lay..there for
about..half an hour.....and finally..got..up.
...We..saw the light of day again...anyway...
didn't..we?

This composition has six pauses of one and a half seconds or longer. Five of these pauses clearly separate groups of words that represent thought units. The one between the words "thinking" and "we" in the twelfth line may be considered doubtful, but it is possible to group the words in the sentence so that the word "thinking" is a unit by itself.

The graph shows how uniformly the first fifty words of the composition are spaced.

Similar records and graphs of the other four compositions with lowest indices of word flow were also made, and all five of these records indicate that on the whole the subjects wrote in thought units. The composition of Subject 26 is an exception, but there is some evidence of confusion in the first part of this composition where practically all of the pauses fall.

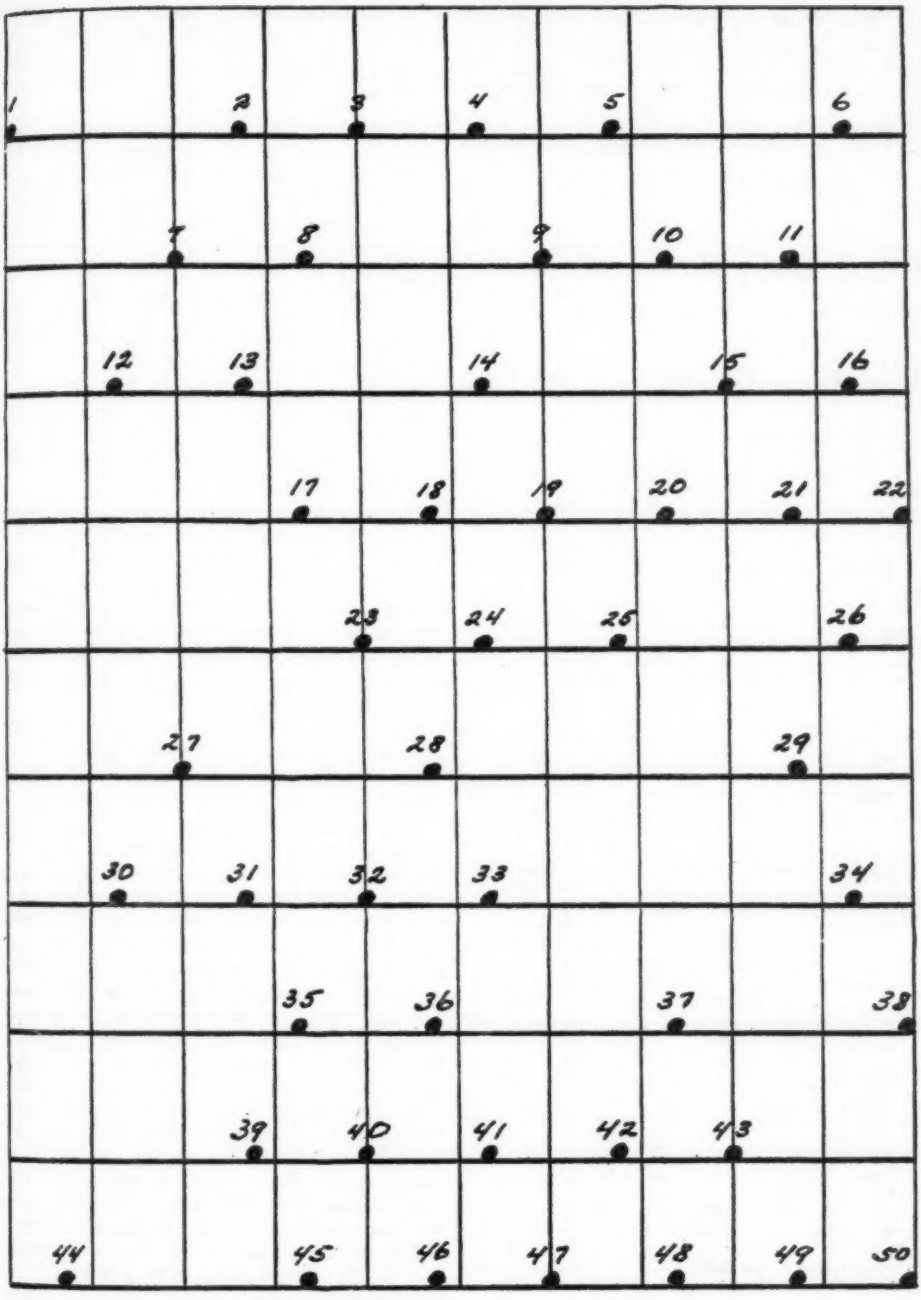


Fig. 2. Rhythm of Word Flow of Subject 79.

Records of the compositions with highest indices of word flow.—The composition with the highest index of word flow, 4.33, was written by Subject 1. The record follows:

1 2 3 4
My.....most...embarrassing.....experience
5 6 7 8 9
....happened...when..I..was.....about
10 11 12 13 14
..ten.....y e a r s...of...old.....My
15 16 17 18 19 20
.....mother....had..just....given..me...a...
21 22 23 24 25
k i t t e n.....for...my...b i r t h...d a y
26
.....T h a t
27 28 29 30 31
....night.....she....was...going...to.....
32 33 34
give..a..party.....
35 36 37 38 39 40 41
..I..wanted....to....stay...up.....but...she
42 43 44 45 46 47 48
.....said....I had...to...go to..bed.....
49 50
.....Of course....I..
didn't.....go....to..s l e e p...but.....just
.....laid.....there.....
..... After....a....while.....the
..... blankets.....started.....
feeling..heavy.....around my....neck so.....
.....I.....reached...up.....
...to.....pull..the.....off..*who*.....
.....*I*.....
.....
....I.....started.....pulling.....
t h e m.....without.....thinking.....
.....
suddenly.....I *her*..heard....my.....cat..let
.....out.....a.....loud....meow,....
.....t h e n
....I.....f e l t.....my.....
hands....around..something.....furry.....
..... Without.....
..... thinking.....I.....
started.....to..*sc*.....scream.....
..... My.....bedroom.....*w*
.....was...right.....next.....to.....
.....the.....living.....room.....
.....a n d..bother...my.....mother.....
and...my..dad..c a m e.....run
.....ning.....in.....
..... My...dad..turned.....on.....the..
light.....a n d....there.
..lay..my..c a t.....right....
..around...my.....n e c k.....
.....A f t e r.....
hitting...the.....cat a few...times.....
and....putting.....h i m....in.....
the.....kitchen.....
my....d a d.....came.....back.....in...
in.... the.....bedroom.....
.....he.....g a v e...m e....a....
good sound....paddling.....
.....just.....to..wake.....
.....me up.
.....Now..every....time.....
.....I see one of....those..ladies.....
.....laugh...at..me.....
This..was..my..most.....embarrassing....
moment.

There are thirty-seven pauses longer than three seconds in this composition. Two of these pauses, those in the nineteenth and twentieth lines, follow crossed out words. Although it is quite evident that the crossing out of these words was due to lack of clarity concerning the order of events in the story, it is difficult to classify the pauses with either those that come at the end of a thought unit or those that break up thought units. They therefore will not be considered.

Of the remaining thirty-five pauses longer than three seconds, twenty-one come at the end of groups of words that express units of thought, while fourteen fall between the words of a thought unit.

There are another thirty-seven pauses from two to three seconds long. One of these can be associated with a crossed out word, but nine occur at the end of thought units and twenty-seven break up units of thought. Inasmuch as the number of pauses that break up thought units increases so rapidly when the pause length is reduced, it is best to consider only those pauses longer than three seconds.

The graph of the first fifty words of this composition indicates a variable rate of word flow. There are five comparatively long pauses, and they come at irregular intervals.

The records and graphs of the other four compositions with high indices of flow of words reveal that those subjects also used long pauses at irregular intervals.

A summary of the analyses made of the records of the five compositions with the lowest indices of word flow and of the records of the five compositions with the highest indices of word flow shows that in the former group there were forty-five long pauses. Eleven of these pauses, or 24 per cent, broke up units of thought. In the five compositions with the highest indices of word flow there were 108 long pauses. Of these, forty-one, or 38 per cent, broke up thought units. This comparison indicates that on the whole the compositions with low indices of word flow are more definitely written in thought units.

The graphs illustrate the fluency of word flow in the compositions. The graphs of the compositions with lowest indices reveal

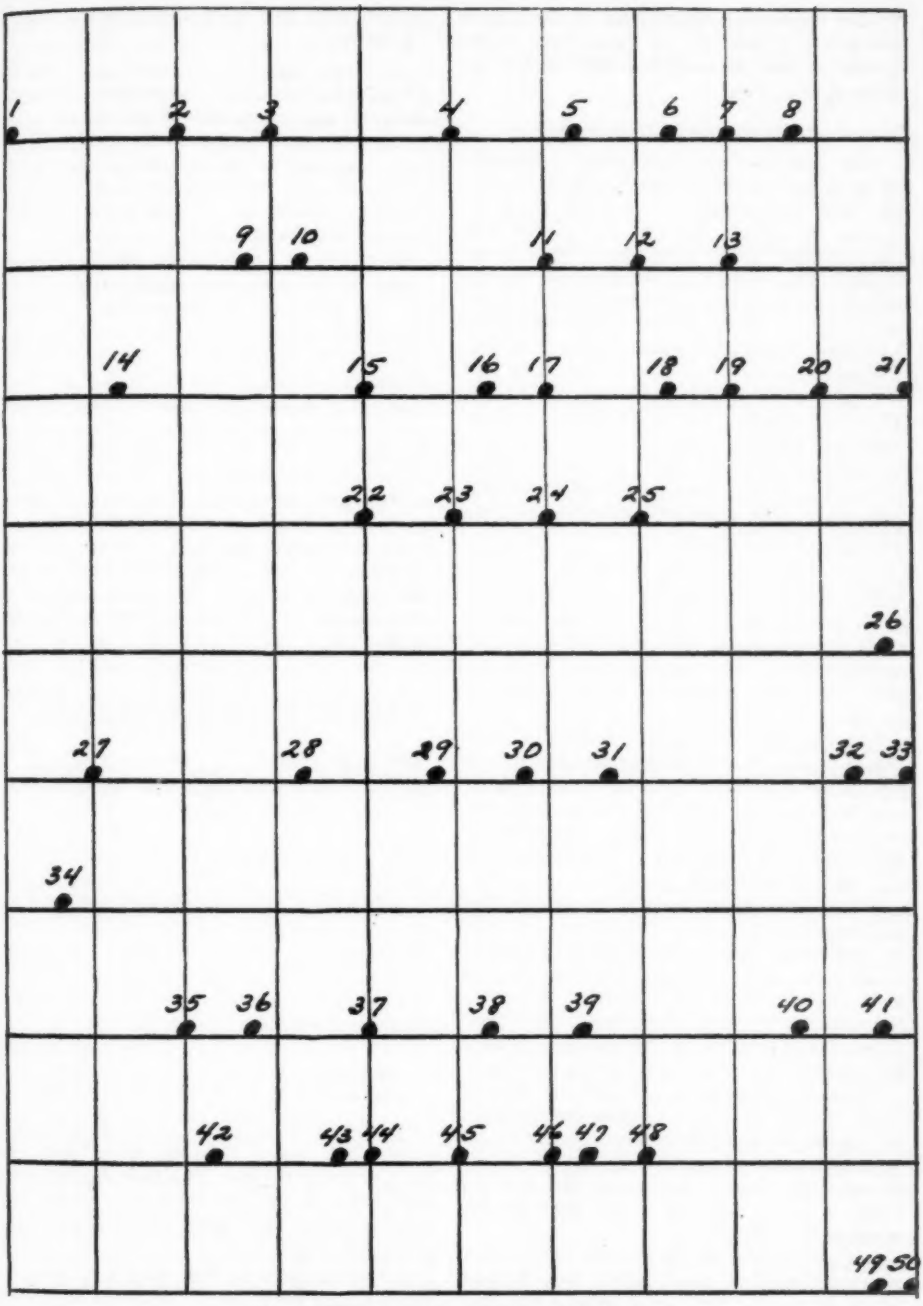


Fig. 3. Rhythm of Word Flow of Subject 1.

greater uniformity and rhythm in the rate at which words come to the writer than do the graphs of the compositions with highest indices of word flow.

Summary and Conclusions

For purposes of comparison a detailed analysis was made of each of the records of the nine compositions that were given the highest rating for quality and of the nine compositions that were given the lowest ratings. The records were compared for the purpose of determining differences in composing structure; that is, the number, length, and location of pauses, between good and poor compositions.

The records of the five compositions with the lowest indices of word flow were also analyzed and compared with the records of the five compositions with the highest indices of word flow. The purpose of these comparisons was to learn if writing in thought units facilitates flow of words.

A summary of the results of the comparisons is given in the following statements:

1. Analysis of the records of the compositions reveals great differences among individual writers. There are differences in length, frequency, and location of pauses between words.

The lengths of the pauses vary with individual pupils. At one extreme are those pupils who pause infrequently during their composition writing; and when they do hesitate, they do so for only a short time. At the other extreme are those subjects who pause frequently and for long periods of time.

Pupils also differ greatly in the number and location of pauses that occur in their compositions. Some group words into thought units while they write and pause only when they come to the end of a thought unit. Others apparently write by words, for their pauses come between the words of a thought unit as well as at the end of thought units.

There are differences even in the flow of words within groups. In some compositions the words in a group follow each other in rhythmic order with short, consistent pauses between the words. In others the flow of words is jerky, and the pauses between the words vary in length.

2. Comparisons of the records of the nine compositions that were given the highest rating for quality and the records of the nine compositions that were rated lowest for quality reveal:

a) That the best compositions, on the whole, were more definitely written in thought units. Of the 139 pauses of significant length in the nine best compositions 115, or 83 per cent, occurred at the end of thought units; and of the 154 pauses in the compositions of poorest quality 107, or 69 per cent, came at the end of thought units.

b) That, on the whole, the pauses are longer in the poor compositions than in the best compositions. Of the pauses longer than two and a half seconds 115 occur in the nine best and 165 occur in the nine poorest compositions.

c) That when the number of crossed-out words or groups of words is used as a measure of certainty of content, there does not appear to be any relationship between compositional quality and certainty of content. Table 3 gives the distribution of the number of crossed-out words for each of the two groups of compositions. It is significant that the two median as well as the two mean number of crossed-out words are approximately the same.

Table III

Number of Crossed-Out Words in the Nine Best and Nine Poorest Compositions

Number of Crossed-out Words or Groups of Words	Best Compositions	Poorest Compositions
12	1	----
11	----	----
10	----	----
9	----	----
8	----	1
7	1	----
6	----	2
5	----	----
4	1	1
3	2	1
2	1	----
1	2	3
0	1	1
Median	2.75	3.00
Mean	3.67	3.33

d) Sufficient evidence to conclude that a relationship exists between word flow and compositional quality. The best compositions are written in thought units, with a rhythmic flow of words, and with comparatively few and short pauses.

3. Comparisons of the records of the five compositions with lowest indices of word flow

and the records of the five compositions with highest indices of word flow reveal:

(a) That there is a relationship between rate of word flow and writing in thought units. Of the 45 significant pauses in the five compositions with lowest indices of word flow 34, or 75 per cent, separated groups of words representing thought units; and of the 108 significant pauses in the five compositions with highest indices of word flow 67, or 62 per cent, came at the end of thought units.

(b) That there appears to be a relationship between rate and uniformity of word flow. On the whole, the flow of words in the five compositions with lowest indices of word flow was rhythmic and uniform; while that of the compositions with highest indices of word flow was jerky and irregular.

(c) That when the number of crossed-out words or groups of words is used as an indication of certainty of content, the evidence points to a positive correlation between rate of word flow and certainty of content. The pupils with the lowest indices of word flow are more sure of their stories than those with the highest indices. Of the former group of five, one had one crossed-out word, one had two, and three had none; of the five with the slowest rate of word flow two had three crossed-out words, one had four, one had five, and one had seven.

4. The number of records examined is relatively small and the results of the comparisons cannot be accepted as highly reliable, but the data indicate:

(a) That good compositions are written predominantly in thought units. Whether writing by thought units is indicative of a linguistic maturity that results in the writing of better compositions has not been determined. Investigations in this area should prove interesting and fruitful.

(b) That writing in thought units accelerates the flow of words. Again, it has not been determined if writing in thought units is indicative of a linguistic maturity that results in a more rapid flow of words. This area also should be investigated.

(c) That slow composing rate is accompanied by much uncertainty during writing. If the number of crossed-out words is an indication of lack of careful planning before beginning the actual writing of the composition, then careful planning helps to accelerate the flow of words.

SECTION VII

SUMMARY AND CONCLUSIONS

Description of the Investigation

The rate of flow of words during composition writing was investigated in its relation to factors within the compositions and academic, environmental, and personal qualities of the writers.

Eighty-four pupils from grades seven, eight, and nine were asked to write two compositions. The first was a reproduction of a story that had been read just prior to the writing and the second was a composition on one of six suggested topics. Records of the time required for the actual writing of each word and of the number and length of pauses between words were obtained by means of a motor driven, time recording kymograph. The average of the larger half of the pauses was taken as the index of the flow of words.

The measure of word flow was correlated and compared with factors within the compositions such as rate of writing, familiar and unfamiliar material, predominance of certain linguistic constructions, length, and quality of compositions; with the academic abilities of the writers in reading, vocabulary, spelling, language, and appreciation of literature; with personal factors including chronological age, mental ability, personality traits, and hand and eye dominance; and with environmental factors such as socio-economic status and place among other children in the family.

A detailed analysis was made of the records of the compositions with highest and lowest ratings for quality and of the compositions with highest and lowest indices of word flow. The results were compared to find the differences in the composing structure of the four types of compositions.

Conclusions

While the small number of subjects used in this investigation tends to limit the reliability of the evidence discovered and while the conclusions must be regarded as merely suggestive, they are valuable in that they point to relationships that have not previously been brought to light and areas that are in need of further research.

An investigation of the time required for the actual writing revealed that when pupils reproduce a selection from memory they write more slowly than they do when writing com-

positions; but in composition writing they write more slowly when they compose the story as they go along than they do when they reproduce a familiar story.

A comparison of writing rate with the rate of flow of words showed that on the whole the rapid writers enjoy a rapid flow of words and the slow writers have a slow rate of word flow. Whether an individual's writing rate decreases or increases with a corresponding decrease or increase in his rate of word flow was not investigated, but the comparisons do suggest that some pupils have a slow reaction time which affects both their writing speed and their flow of ideas.

A comparison of the rate of flow of words in the two compositions revealed that the index of word flow is remarkably consistent. The coefficient of correlations between the indices of the first group of compositions and those of the second group was .77 with a probable error of .03. In writing the second group of compositions, which required composing as they went along, the pupils showed a slower rate of word flow than they did in the first group of compositions.

Linguistic constructions within the compositions were investigated in their relation to the rate of flow of words. Although no relationship could be found between the Adjective-Verb Quotient and flow of words, it was discovered that compositions with very low indices of Subordination and high Personal Pronoun Indices are written with a very slow flow of words, but that those with high Indices of Subordination and low Personal Pronoun Indices are also written with a slow flow of words, though not as slow as that of the first group. Inasmuch as the evidence is not conclusive as to what these linguistic constructions definitely determine, it is difficult to know what their relationship to the flow of words signifies.

On the whole the shortest compositions as well as those of poorest quality were written with the slowest flow of words, and the longest compositions showed the most rapid rate of word flow. The compositions that were ranked highest in quality, however, did not have the best rate of flow of words.

A study of the relation of academic factors to rate of word flow revealed that the pupils with the lowest scores in reading, vocabulary, and spelling have the slowest flow of words. The coefficients of correlation between these subjects were .85 for reading and vocabulary, .65 for reading and spelling, and .68 for vocabulary and spelling. These relatively high coefficients suggest that there may be some common element in these subjects that affects the rate of word flow. More research in this area may prove fruitful.

The comparisons also revealed a similarity in the relation of English marks and scores in language usage to the rate of word flow. The coefficient of correlation between these two academic measures is .48, but their relation to rate of word flow is almost identical, especially for the second group of compositions. A likely explanation is that teachers, in marking the ability of the pupils in English, place a great deal of emphasis on language usage.

A study of the relationship between scores in prose appreciation and rate of word flow revealed that pupils with high scores in prose appreciation are inclined to pause a great deal during composition writing.

A coefficient of correlation of .69 between the scores in Mechanics and Effectiveness of Expression indicates that they measure related aspects of English ability; but a comparison of the scores in these subjects with the indices of flow of words revealed that this particular aspect of English ability has no relation to the rate of word flow.

In considering the personal traits of the subjects, it was learned that rate of flow of words increases with increase in chronological age, mental age, and in rise from grade to grade in school. The possibility that these factors may be merely subsidiary and that the increase in rate is due primarily to the additional practice in composition writing which is provided during the junior high school years has not been studied. Further research is needed to determine whether increased rate of flow of words is an accompaniment of increase in chronological age and mental age or the result of additional practice.

The intelligence quotient, although a measure of mental ability as well as mental age, has a different effect on the rate of flow of words. The relationship is curvilinear, with a rapid rate of word flow shown by pupils with both high and low intelligence quotients.

A relationship was also found between rate of flow of words and some of the personality traits. A high degree of perseverance seems to be reflected in a rapid flow of words, and a

decrease in perseverance appears to be accompanied by a decrease in rate of word flow, but pupils with a low degree of perseverance tend again to have a rapid flow of words. The relatively high coefficient of correlation between intelligence quotients and perseverance ratings, .59, gives rise to some speculation as to the relationship between the two. Their relationship to rate of word flow is similar, and the influence of one upon the other is difficult to determine.

Pupils who by their scores on personality tests indicate that they are dominating, extrovertive, and emotionally stable, enjoy a rapid flow of words. Those who are introvertive and emotionally unstable have a slow word flow.

A comparison of the composing rates of girls and boys revealed that while both sexes have approximately the same rate in the seventh grade, the girls increase their rate more rapidly than the boys as they proceed through the grades.

A study of hand and eye dominance in its relation to rate of flow of words revealed that pupils with right hand and right eye dominance, on the whole, compose with fewer and shorter pauses than those with left hand and left eye dominance. Those with crossed hand and eye dominance have an extremely slow flow of words. The matter of dominance, especially of crossed dominance, is in need of further investigation.

Since the average chronological age of the pupils with right hand, right eye, and right hand-right eye dominance is 163 months compared with 158 months for those with left hand, left eye, and left hand-left eye dominance, it is difficult to determine to what extent the difference in rate of flow of words is due to the difference in chronological age. The sixteen pupils with crossed hand-eye dominance had a mean chronological age of 160 months. Their mean index of word flow is considerably higher than that of any of the other groups. Inasmuch as the mean chronological age is approximately half way between the others and the mean index of word flow is higher, it seems reasonable to conclude that the pupils with crossed hand-eye dominance have a slower rate of word flow. The number of cases is too limited, however, to accept this conclusion as completely reliable.

A comparison of the mean indices of word flow for the five levels of socio-economic status revealed that with a rise from Level II to Level IV there is a decrease in rate of flow of words. The mean chronological age for these three groups ranges from 162 to 164 months; so the change in rate cannot be attributed to the influence of chronological age. The evidence suggests that pupils in the higher levels of socio-economic status have a better rate of word flow than those in the lower levels.

The mean index of word flow for Level I, however, rises sharply while the mean index for Level V shows a sharp decline. The fact that the mean chronological age is 158 months for the pupils in Level I may account for the rise in the mean index, but the mean chronological age for the pupils of Level V is 162 months and cannot be claimed to affect the rate of word flow. The number of cases in these two groups, however, is extremely small and very likely accounts for the change in trend of the mean indices more than any other factor.

The comparison of the mean indices of word flow of only children, oldest children, youngest children, and others revealed an extremely small difference in rate of word flow of these four groups. Evidently, a child's place among other children in the family does not affect his rate of word flow.

A detailed analysis of the records of the nine best and nine poorest compositions indicated great differences among individual pupils in pattern of word flow. Not only is there a great difference in the number and lengths of the significant pauses between groups of word of a thought unit, but the flow of words within groups also varies.

Comparisons of the records revealed that the best compositions are written in thought units, while the poorest compositions give evidence of having been written by words. The poorest compositions are also written with longer pauses.

The records of the five compositions with lowest and the five compositions with highest indices of word flow indicate that writing by words is accompanied by a slow flow of words and writing by thought units is correlated with a more rapid word flow. The compositions written with a slow composing rate manifest very little rhythm in the flow of words.

Educational Implications

The educational implications of the results of this investigation are numerous. A few of them are:

1. The findings of this study corroborate the claim that English is closely related to other school subjects. Pupils who were poor in reading, vocabulary, and spelling indicated a tendency to pause often when writing compositions. English cannot be taught in isolation from other subjects.

2. The teaching of the mechanics of writing does not necessarily result in more fluent composing. Ability in mechanics and effectiveness of expression did not seem to be associated with rate of flow of words.

3. The problem of personality development cannot be divorced from teaching. The dominating, extrovertive, and emotionally stable pupils wrote with a rapid flow of words while those with introvertive tendencies and lack of emotional balance paused often and long during composition writing. It is evident that something must be done for the latter group to give them more confidence and place them more at ease if they are to use their abilities to the best advantage and show improvement in composing rate.

4. Pupils should be taught to plan their writing. The rate of flow of words was much better when pupils wrote familiar material than when they wrote without much preparation. The compositions with the slowest rate of word flow also had the largest number of crossed-out words, and they gave evidence of having been written by words rather than by thought units. Pupils should be encouraged to think through their stories before writing them. There is evidence to believe that this will lead to the writing by thought units and to an increase in rate of word flow.

5. Mere composing fluency should not be the primary objective in composition teaching. Pupils with the lowest scores in perseverance, English, and language usage had a rapid flow of words. Composing fluency that is associated with and the accompaniment of carelessness and lack of perseverance should be discouraged.

On the other hand, the best spellers and those with the highest scores in prose appreciation had a comparatively slow flow of words. Composition teaching should, therefore, encourage a flow of words that is not so rapid as to promote carelessness but rapid enough to stimulate continuity both of thinking and of expression.

6. Inasmuch as there appears to be a positive correlation between writing by thought units and both rate of word flow and quality of composition, composition teaching should emphasize thinking and writing in thought units.

Bibliography

Allport, G. W., and Cantril, H. "Judging Personality from Voice," *Journal of Social Psychology*, V (February, 1934), 37–56.

Allport, Floyd H., Walker, Lynnette, and Lathers, Eleanor. "Written Composition and Characteristics of Personality," *Archives of Psychology*, No. 173, XXVI (December, 1934), pp. 82.

Anderson, John E. "An Evaluation of Various Indices of Linguistic Development," *Child Development*, VIII (March, 1937), 62–68.

Bear, Mata V. "Children's Growth in the Use of Written Language," *Elementary English Review*, XVI (December, 1939), 312–319.

Boder, David P. "The Adjective-Verb Quotient: A Contribution to the Psychology of Language," *The Psychological Record*, III (March, 1940), 309–43.

Eisenson, Jon. "Some Characteristics of the Written Speech of Stutterers," *Pedagogical Seminary*, L (June, 1937), 456–58.

Healy, Katharine L. "A Study of the Factors Involved in the Rating of Pupils' Compositions," *Journal of Experimental Education*, IV (September, 1935), 50–53.

Hoppes, William Calvin. "The Development of Written Expression among Children of Elementary School Grades," Unpublished Ph. D. dissertation, Department of Education, University of Chicago, 1931.

La Brant, Lou L. "A Study of Certain Language Developments of Children in Grades Four to Twelve, Inclusive," *Genetic Psychology Monographs*, XIV (November, 1933), 387–492.

La Brant, Lou L. "The Occurrence of Dependent Clauses in the Writing of School Children in Grades Four to Nine," *Educational Trends*, I (January, 1932), 16–20.

Language in General Education, A Report of the Committee on the Function of English in General Education for the Commission on Secondary School Curriculum (New York: D. Appleton–Century Co., 1940), pp. xii + 226.

Lowen, Maud B. "The Effect of Environment on Creative Ability in Fourth-Grade Children," *Elementary School Journal*, XXXVI (October, 1935), 120–26.

Lyman, R. L. "Fluency, Accuracy, and General Excellence in English Composition," *School Review*, XXVI (February, 1918), 85–100.

Sanford, Fillmore H. "Speech and Personality: A Comparative Case Study," *Character and Personality*, X (March, 1942), 169–98.

Shepherd, Edith Elizabeth. "The Influence of Content and Maturity on Sentence Structure in Pupils' Written Discourse." Unpublished Master's thesis, Department of Education, University of Chicago, 1933.

Sims, Verner Martin. *The Measurement of Socio-Economic Status* (Bloomington, Illinois: Public School Publishing Co., 1928), pp. 33.

Stagner, Ross. "Judgments of Voice and Personality," *Journal of Educational Psychology*, XXVII (April, 1936), 272–77.

Taylor, Harold C. "Social Agreement on Personality Traits as Judged from Speech," *Journal of Social Psychology*, V (May, 1934), 244–48.

Teaching Language in the Elementary School. Forty-third Yearbook of the National Society for the Study of Education, Part II (Chicago: National Society for the Study of Education, 1944), pp. ix + 257.

A STUDY OF CHILDREN'S HEROES AND IDEALS

M. Louise Stoughton
Samuel Bowles School
and
Alice M. Ray
Brightwood School
Springfield, Massachusetts

This study was made with 344 children in the second, fourth, and sixth grades in three schools in Springfield, Massachusetts.[1]

Procedures

It can be noted that the procedures in this study have been employed in numerous earlier investigations. One of the more recent studies, comparable to the present investigation but involving a considerably larger number of children, is a study by Hill.[2]

Two questions were presented to the children: Of all the persons whom you have known, or heard about, or read about, whom would you most wish to be like? And why do you like this person?

The responses were obtained from the second-grade children through individual interviews. In the fourth and sixth grades the children wrote their responses on a form that contained the above two questions. Neither the teachers nor the children had any knowledge of the questions before the papers were distributed. Precautions were taken to avoid any suggestions by teachers or classmates which might have influenced the children's choices.

Subjects

The 344 children included 170 boys and 174 girls. Springfield, Massachusetts, the city in which this study was conducted, has a population of about 160,000. It is a city of finance, commerce, and industry, manufacturing a diversified line of products and employing a large number of skilled workers.

The schools were situated in cosmopolitan neighborhoods where the families lived in average economic circumstances. A number of nationalities were represented.

[1] This study was conducted with the advice of Professor Arthur T. Jersild, Teachers College, Columbia University.

[2] D. S. Hill, "Personification of Ideals by Urban Children." *Journal of Social Psychology*, I (1930), pp. 379–393.

Treatment of Data

The responses given by the children were tabulated under broad categories as follows:

I. Persons or characters in the child's immediate, everyday environment.
   1. Parents.
   2. Relatives.
   3. Friends.
   4. Teachers.
   5. Others in the community.

II. Persons or characters remote from the child's everyday environment, including fictional and imaginative characters.
   1. Characters from literature.
   2. Characters from history.
   3. Characters from the news.
   4. Characters from the comics, movies, and radio.

III. Religious characters.

Findings

The main results of this study are shown in Table I. It will be noted that in the majority of cases ideals were chosen from the immediate environment (51%). In Hill's study more than half (57%) were chosen from the remote environment. This difference is probably due in part to the fact that Hill included young people up to twenty years of age, while in this study the oldest children were only thirteen years old.

Analysis of the responses showed that characters from the remote environment were named more often by the boys than by the girls, 65% of the boys choosing as ideals characters from that environment, as compared with 30% of the girls. This trend agrees with Hill's findings.

Only two percent of all the children named a religious character as their ideal.

TABLE I

PERCENTAGES OF HEROES AND IDEALS NAMED IN VARIOUS CATEGORIES BY BOYS AND GIRLS AND BY BOYS AND GIRLS COMBINED AT VARIOUS GRADE LEVELS

	Grade 2			Grade 4			Grade 6			All	All	All boys &
	Boys	Girls	All	Boys	Girls	All	Boys	Girls	All	Boys	Girls	girls
No.	41	44	85	61	55	117	67	75	142	170	174	344
Categories												
I Persons or Characters from Immediate Environment	58%	84%	72%	31%	65%	47%	21%	60%	42%	33%	68%	51%
1 Parents	7.3	15.9	12	9.7	9.1	9	4.5	9.3	7	7	11	9
2 Relatives	22.0	13.6	18	3.2	3.6	3	4.5	1.3	3	8	5	7
3 Friends	24.4	26.8	26	6.4	23.6	15	3.0	16.0	10	10	21	15
4 Teachers	0	27.3	14	3.2	20.0	11	6.0	24.0	16	3	24	14
5 Others	2.4	2.3	2	8.1	9.1	9	3.0	9.3	6	5	7	6
II Characters from Remote Environment	37	14	25	69	33	52	78	39	57	65	30	47
1 Literature, fiction	4.9	4.5	5	12.9	1.8	7	6.0	16.0	11	8	8	8
2 History	4.9	4.5	5	21.0	14.6	18	13.4	10.6	12	14	10	12
3 News, Contemporary affairs	2.4	0	1	11.3	0	6	43.3	4.0	23	22	2	12
4 Comics, Movies, Radio	24.4	4.5	14	24.2	16.4	21	14.9	8.0	11	21	10	15
III Religious Characters	5	2	3	0	2	1	1	1	1	2	2	2

* "Parents" represents mention of mother by 5.5% and of father by 3.4% of the children.

The following percentages taken from Table I show that as the children grew older there was a diminishing interest in the immediate environment and an increasing interest in the remote.

	Grade 2	Grade 4	Grade 6
Immediate Environment	72%	47%	42%
Remote Environment	25%	52%	57%

Analysis of choices from the immediate environment showed that among the boys, friends (10%) were most frequently idealized, followed by relatives (8%). Among the girls, teachers (24%) were most often chosen as ideals, with friends (21%) coming next. The boys' choice of friends and relatives decreased and their choice of teachers increased from the second to the sixth grade. The girls' choice of teachers remained about the same in all three grades. Parents were idealized more by second-grade children than by older children.

Analysis of choices representing the remote environment shows that 22% of all the boys and only two percent of all the girls chose as ideals persons who were prominent in the news. Contemporary characters were especially popular with the sixth-grade boys, 43% indicating this interest.

Ideals from literature were chosen about equally by boys and girls. As would be expected, older children made more choices from literature than did younger children.

Historical characters were the most popular among the fourth-grade children. This was undoubtedly the result of the fourth-grade social studies which deals with exploration and colonization.

Characters in movies, comics, and radio were more often named as ideals by fourth-grade children than by either older or younger children. This group of characters represents the ideals chosen by 21% of all the boys, which places it next in favor to characters in the news and public life.

Table II shows the extent to which boys and girls chose masculine and feminine characters, regardless of classification.

TABLE II

	Grade 2		Grade 4		Grade 6		All Grades	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Masculine Characters	85%	11%	95%	11%	94%	7%	92%	9%
Feminine Characters	10%	86%	5%	87%	5%	92%	6%	89%

Throughout the grades masculine characters were more popular among the boys, and feminine characters among the girls.

It was interesting to note that there were more girls who chose masculine characters for ideals than there were boys who chose feminine characters. Feminine characters were chosen by second-grade boys more than by the older boys.

IDENTITY OF CHARACTERS OR PERSONS NAMED IN VARIOUS CATEGORIES

Parents.—The following table shows the number of children in each grade who chose their fathers or mothers.

	Grade 2	Grade 4	Grade 6	All Children
Father	3	7	2	12 (3.4%)
Mother	7	4	8	19 (5.5%)

The fathers were chosen by boys and the mothers by girls in all but two instances. One sixth-grade boy named his mother and one fourth-grade girl, her father.

Relatives.—Among the relatives, uncles received the highest number of choices. This might have been due to the fact that several of the uncles were in military service. Next in order came sisters, brothers, cousins, and aunts. Children in the second grade chose relatives more often than did the children in the upper grades.

Friends.—Friends were chosen by second-grade children more than by the older children. Playmates seemed to loom large in the eyes of the younger children.

Teachers.—The teachers named as ideals were classroom teachers, principals, and supervisors of special subjects.

Others in the community.—This group included Cub Scout leaders, Scoutmasters, neighbors, dancing and piano teachers, a doctor, and a librarian.

Literature.—*Caddie Woodlawn*, *Robin Hood*, and *Charles Atlas* were each named twice by sixth-grade children, and *Tom Sawyer* and *Uncle Ben* twice by fourth-grade children. Some of the characters named only once by children in these grades were *Huckleberry Finn, Peter Rabbit, Dave Dawson, Nan Bobbsey, Homer Price, Judy Bolton, Nancy Drew, Sue Barton*, and *Cherry Ames*. Only two characters from literature were named by second-grade children,—*Santa Claus* and *Mother Bear*.

The following lists give the names of characters and the number of times they were chosen in each grade:

History.—

	Grade 2	Grade 4	Grade 6	All
Clara Barton		4	5	9
George Washington	3	2	2	7
Abraham Lincoln		1	3	4
Daniel Boone			2	2
Stephen Foster		2		2
Christopher Columbus		2		2

Other characters named only once were *Betsy Ross, Thomas Edison, Champlain, Florence Nightingale, Eli Whitney, Robert Fulton*, and *Alexander Graham Bell*.

News.—

	Grade 2	Grade 4	Grade 6	Total
Pres. Roosevelt	1	1	6	8
General MacArthur		3	6	9
Major Bong		1	2	3

Included among the characters named once were *Daniel Beard Carter, Gen. Patton, Charles Goodyear, Joe Dimaggio, J. Edgar Hoover, Raymond Gallagher* (Chief of

Police), *Jack Dempsey, Babe Ruth, Mrs. Eleanor Roosevelt, Mrs. Mary Bethune, Richard E. Byrd,* and *Lord Louis Mountbatten.*

Movies, comics, and radio.—

	Grade 2	Grade 4	Grade 6	Total
Gene Autry	3	1	2	6
Roy Rogers		4	1	5
Margaret O'Brien	1	3		4
Shirley Temple		2	1	3
Bing Crosby	1	2		3
Lone Ranger		1	1	2
Dick Tracey		2	1	3
Jack Benny	1	1		2
Bob Hope	1	1		2
Superman	1	1	1	3

Other characters chosen once were *Van Johnson, Rochester, Frank Sinatra, Deanna Durbin, Dorothy Lamour, Dr. Wassell, Phil Baker, Elizabeth Taylor, Sonja Heine,* and *Flip Corkin.*

Factors Influencing Choices

The indefinite quality of "goodness" was one of the outstanding characteristics that appealed to the children of all grades.

In 17% of the cases characters were chosen by the second graders because they "gave them things." This factor was mentioned considerably less in the upper grades.

It was amusing and enlightening to note that no second- or fourth-grade child considered his teacher an ideal because of her intelligence. That distinction was reserved for the principal of the school!

Another interesting observation was the fact that some children admired in others that which was lacking in their own make-up, e.g., the little second grader with bobbed hair who wanted to be her girl friend because she had long braids; or the little girl who liked a classmate because "she is a good sport, she is nice, she is kind, and she is a clean girl."

Adventure and courage were the traits that seemed to make the greatest appeal to the fourth-grade boys. Although only ten actually gave adventure as the reason for their choice, they implied it by the type of answer they gave, as when one boy chose Tom Sawyer because, "He sails a big boat. He is a pirete. He gos to Ilands at night!"

It was also interesting to note that the only characters chosen from the comics were characters that also appear on the radio, Superman, Flip Corkin, and Dick Tracey.

Some of the children were vague and indefinite in the reasons given for their choice of heroes; others, however, seemed to have a clear idea as to the qualities they admired. These ranged all the way from the response of a second-grade boy who wished he might be Mother Bear because "She always had porridge," to the reason given by a boy who wanted to be like Huckleberry Finn because "He had a lot of adventures, a lot of fun, and lived in what I would call a paridise!"

Some of the responses were interesting when considered in relation to the circumstances of individual children.

A girl who was one of ten children named as her ideal a playmate who was an only child "because I have fun playing with her and she doesn't have any little brothers and sisters to take care of."

Another girl chose a playmate whom she liked because "she is not too smart and not too dum, not too rich and not too poor, because she is just an average girl."

A small girl wished to be like "any common second grade teacher."

One girl wished to be like the teacher of physical education because, "I like the way she teaches. It seems to be very easy to learn."

A boy named a teacher of arithmetic because, "I would like to teach what he teaches and have the answers ready for all the questions I am asked."

An original spelling of "education" appeared when a child admired his principal because she had a good "ageakshun."

General Comments

The results of this study show that children's interests broaden as they grow older. A majority of children in the second grade

found ideals through personal experience in the immediate environment: home, school, and neighborhood. The older children more often found ideals in the larger world as represented by history, the movies, and contemporary life and events. Interest had extended beyond the local community to include broader world concepts. Growth of reading ability undoubtedly played a part in opening the way to wide interests.

In certain cases there were glimpses of maladjustment of individual children. Some responses suggested the need of psychological guidance.

This inquiry into the ideals of children has been very revealing to the investigators. The writers, in the process of conducting this study, became convinced that knowledge of the qualities children admire in others would help teachers do more effective work with children.

CHILDREN'S ATTITUDE TOWARD DISCIPLINE AS RELATED TO SOCIOECONOMIC STATUS

LAURA DOLGER and JANET GINANDES[1]

The object of this study was to investigate the attitudes toward discipline exhibited by children representing different socioeconomic and cultural groups. To achieve this object a segment of the school population was studied in two schools in the same city (New York) which differed widely in their educational philosophy and in the economic and social status of the pupils.

SUBJECTS

We shall hereafter designate the schools by their initials. School P. S. is a conservative public school located in mid-town Manhattan. Its school population includes children of Puerto Rican, Irish, Italian and Negro stock. Many children are of foreign born parentage who migrated to this country. Many of the mothers worked to supplement the family income and the children frequently assumed part of the household responsibilities. Most of the families had multiple siblings, usually more than two. Some of the families at times relied upon home relief for their support.

In contrast to school P. S., school D. is a progressive private school located a short distance away. The children come from families of the middle and upper middle class income groups. A few of the children attend the school through scholarships. By and large, finance is not apparently a disturbing problem to these children and their families. The family unit is generally small (two children or one) and many of the children were, in their earlier years, cared for by nurses or governesses. Many of these children were given allowances for their own free personal use.

Twenty-one children were selected from each of the two schools for the purpose of this study. The selection was made at random, the only requisites being that the children were in the fifth grade and averaged between nine and a half and ten and a half years. The groups included boys and girls.

[1] This study was undertaken in connection with a course conducted by Dr. Arthur T. Jersild and Dr. Charlotte Fehlman.

METHODS

The procedures used in the study follow:

1. The children were asked to write a composition based on a story which was read to them in class by their teachers. The story was a simple tale of two children:

> A younger child, Jimmy, aged six, constantly interfered with his older brother, Tommy, aged ten. Whenever Tommy entertained his friends, Jimmy intruded; when Tommy wanted to play, Jimmy took his toys. Jimmy always broke up Tommy's games and was generally considered a poor sport.

After this story had been presented each child was asked to write in a composition what he would do about Jimmy if he were Tommy. Through this medium the child was expected to indicate his disciplinary attitudes in a situation that he himself might experience at any time. Children who had no siblings were asked to imagine the situation.

The compositions of the two groups were collected and the results tabulated with a view to making a separate tally of compositions which (1) offered what appeared to be a "constructive" amicable solution as contrasted with (2) a *"non-constructive"* disciplinary solution with an appeal to authority. Some interesting notes were also made on the physical appearance of the compositions. The data obtained and relevant comments will be made at a later point.

2. In addition, each child was interviewed. Each interview lasted about twenty minutes. A series of ten questions relating to routine disciplinary factors common in the school experience were presented. An attempt was made to put the child at ease. He was told that he was not being given a test, that no grading was being done and that he could speak freely. The physical conditions surrounding the interview in the two schools were unavoidably somewhat different. In P. S. the rooms were isolated spare offices. There was quiet and the subject's attention was not

deflected by extraneous noise. In school D. the children were interviewed in one of the classrooms, or in a library or locker room since these places were the only ones available. There was continuous activity about the examiner and child throughout the interview. The attention given the examiner, however, was satisfactory and there seems to be no reason to think that the rooms had any influence on the responses of the pupils since a particular effort was made by the interviewers to put the children at their ease.

The specific questions included in the interview will be summarized in the discussion of results below.

Findings

Compositions.—As indicated above, the solutions offered by the children in their compositions were classified into two groups: the *constructive amicable solution* and the *non-constructive appeal to authority*. By a *constructive amicable solution* is meant any attempt on the part of the child to settle the differences between Jimmy and Tommy on the basis of reason. For example, responses such as, "Teach Jimmy to play fair and learn the rules of the game." or "Invite Jimmy's friends over." or "Give Jimmy his own toys" were treated as representing an effort to resolve the conflict without the use of authority or punitive methods.

This is in contrast to the children who suggested that Jimmy be punished either by the older child or by one in authority. Responses of this character were classified as representing a *non-constructive appeal to authority*.

In the treatment of the data, a separate tally was given for each solution offered by children who happened to give more than one. The findings in this part of the study follow:

	School D (Private, "progressive," pupils of relatively high socioeconomic status)	School P. S. (Public, "traditional," pupils of relatively low socioeconomic status)
No. of children	21	21
No. of solutions offered	38	35
No. of Constructive, amicable solutions	29	18
No. of Non-constructive appeal to authority solutions	9	17

The substance of representative answers in the "constructive, amicable" category follows:

Teach the child how to make something new out of broken toys.
Teach the offending brother the rules and how to play fair.
Give him his own game or toy.
Let the injured party play at friend's house.
Invite the offending brother's friends over to the house.
Make him feel important by letting him make the rules.
Teach him sportsmanship by taking him to professional games.
Reason with him.

Characteristic answer of the non-constructive variety included:

Dictatorial verbal discipline: i.e. scolding.
Recommendations that the offending child be punished by his brother or a parent.
Pay no attention to the child, let him cry.
Retaliation—do the same thing to him.

It can be noted that the children representing a relatively low socioeconomic level gave responses involving a non-constructive appeal to authority as did the children in the other school. The solution offered by a larger proportion of the children in the former group involved an effort to put the offending brother in his place by employing some form of punishment. About three-fourths of the solutions offered by children in school D. were constructive as compared with about a half of the solutions offered by children in school P. S.

The over all picture shows a total of forty-seven tallies in the constructive solutions group and only a total of twenty-six tallies in the non-constructive solutions group. This suggests that those who subscribe to solutions of constructive sort are more resourceful and imaginative than those who use the formulas represented by the non-constructive solutions.

Interview Data

The following gives the ten situations and questions presented during interviews, the children's answers and comments concerning these.

Item 1. Fred's parents could not afford to buy him any toys. Bob had lots of marbles which his parents gave him for Christmas. One day Fred took a few marbles out of Bob's

desk. When the teacher asked who took them, Fred admitted that he had. What would you do if you were the teacher?

The solutions to this question when tallied in terms of the categories already described above were as follows:

Non-constructive solutions: School D.: 1.
School P.S.: 16.

Characteristic answers in this group included:

Send Fred to the principal.
Send a letter to his mother.
Fred's parents should deprive him of dessert.
Fred should be made to stay in after school.

Constructive solutions: School D.: 34.
School P.S.: 14.

Characteristic answers in this group included:

Children should share toys.
Fred should earn money to buy marbles.
Fred should not be blamed because he admitted guilt.
Children or teacher should give Fred a toy.

It can be seen that again a predominantly greater number of non-constructive responses came from school P. S. and a predominantly greater number of constructive responses came from school D. The group which favored constructive solutions were of the opinion that the conditions causing the offense were to be remedied. Those answering in favor of non-constructive solutions seemed to express the view that the guilt was in the child while those favoring a constructive solution seemed to hold the view that the situation was responsible.

Item 2. Should children ever talk back to their parents?

The answers given to this question were as follows:

No (definitely)	School D.:	4
	School P.S.:	15
No (conditionally)	School D.:	14
	School P.S.:	6
Yes	School D.:	1
	School P.S.:	0
Don't know	School D.:	2
	School P.S.:	0

The figures reveal that a greater number of P. S. children were opposed to talking back under any circumstances while the greater number of children of school D. felt that certain conditions justified the children stating an opposing point of view.

Item 3. If there is a class discussion, is it all right to interrupt your neighbor if you have something to say?

The answers to this question showed no appreciable difference between the two schools. Children in both groups apparently believed that it is wrong to interrupt.

Item 4. Is it ever all right to tell a lie?

The answers given to this question were:

No (without exception)	School D.:	4
	School P.S.:	14
No (with exception)	School D.:	17
	School P.S.:	7

Both groups tended to be strongly opposed to lying, but the children in school D. were inclined to name certain instances, such as helping others, when lying was permissible.

Item 5. Should children ever talk back to their teacher?

The answers given to this question were:

No (without exception)	School D.:	6
	School P.S.:	17
No (with exception)	School D.:	15
	School P.S.:	4

It can be noted that, as in Item 2, more children in school D. felt that certain conditions justified the child's stating his point of view.

Item 6. What should the teacher do if a child whispers in class?

The answers to this question fall into two groups: (1) Non-constructive solutions, employing the use of punishment by segregation, deprivation or chastisement and (2) Constructive solutions.

(1) Non-constructive	School D.:	16
	School P.S.:	26
(2) Constructive	School D.:	5
	School P.S.:	2

The answers to this question indicate that a majority of the children in both schools felt that punishment was necessary for this behavior. However, fifteen of the sixteen non-constructive responses from school D. suggested punishment by segregation (e.g. moving the child to another seat) while twenty of the twenty-six non-constructive responses from school P. S. suggested a more severe punishment, such as sending the child to the

principal, telling the parents, or keeping the child after school.

Item 7. If the boy sitting next to you copies your paper during an exam, should you report him? What should the teacher do?

The answers to the first part of this question were:

No	School D.:	11
	School P.S.:	2
Yes	School D.:	9
	School P.S.:	18
Avoid issue by moving	School D.:	5
	School P.S.:	0
Explain why child should not copy	School D.:	2
	School P.S.:	0

There were more children in school D. who were against "tattling" than there were in school P. S.

The answers to the second part of this question fall into two groups: (1) Non-constructive solutions and (2) Constructive solutions.

Non-constructive solutions: School D.: 18.
School P.S.: 18.

Characteristic responses in this group were:

Give test in another room.
Write note to parents.
Stay after school.

Constructive solutions: School D.: 3.
School P.S.: 1.

Characteristic responses in this group were:

Explain to child he will not learn that way.
Explain that wrong answers may be copied.

The majority of children in both schools felt that punishment was necessary in this case. However, thirteen of the eighteen non-constructive responses offered by the school D. children suggested lenient handling such as segregation, while ten of the eighteen non-constructive responses offered by the school P. S. children were more severe, suggesting such action as writing to the parents.

Item 8. What should a teacher do if she finds a child taking money out of her desk?

The answers in response to this question fall into two groups: (1) Non-constructive authoritarian solutions and (2) Constructive solutions.

Non-constructive solutions: School D.: 15.
School P.S.: 37.

Characteristic responses were:

Tell parents.
Spank.
Deprive of play time.

Constructive solutions: School D.: 18.
School P.S.: 7.

Characteristic responses were:

Explain why child should not steal.
Work for money.

In analyzing the data elicited by this question, it can be seen that more of the non-constructive solutions were offered by the school P. S. children and more of the constructive solutions were offered by the school D. children.

Item 9. What should the teacher do if a child plays hookey from school?

The responses to this question again were tallied as non-constructive or constructive responses.

Non-constructive: School D.: 18.
School P.S.: 37.

Characteristic answers in this group included:

Tell mother, principal or truant officer.
Keep after school.

Constructive: School D.: 20.
School P.S.: 6.

Characteristic answers in this group included:

Change the child's school if he does not feel happy there.
Explain the need for attending school.

The data indicate that more of the non-constructive solutions were offered by the school P. S. children and more of the constructive solutions were offered by the school D. children.

Item 10. Is it all right for a teacher to leave her class on their honor?

The responses to this question were:

Yes	School D.:	13
	School P.S.:	0
No	School D.:	2
	School P.S.:	16
Yes (conditionally)	School D.:	7
	School P.S.:	6

The school D. children were overwhelmingly in favor of the honor system as an educational policy. They approved of it as a

means of learning self-control and of building self-confidence and self-reliance. The school P. S. children, on the other hand, expressed their disapproval of the honor system and indicated that most children cannot be trusted when left alone.

Other Observations

The format of the two sets of compositions presented interesting differences. The P. S. papers were all written on white ruled composition paper of the same size. On the top of the page each child had painstakingly written his name, school, grade and the date. The margins were uniform and the children were careful to observe them throughout.

The ideas reflecting the children's thinking were as constrained as the boundaries of their compositions. There was little originality in their thinking; vocabulary was limited and there was a sameness about each paper that was striking. In contrast to this, the D. school children's papers were of every possible size of note book paper. Some of the ideas presented were topically paragraphed, others rambled a good deal, but most of them revealed spontaneity and imagination. There was no attempt at formality, each child writing as he pleased. A greater facility of expression and an independence of thinking was evident.

The interviews also revealed certain additional differences between the two groups. Although an attempt was made to assure the children that they were not being given a test, they nevertheless did not react similarly to this assurance. The D. children were very much interested in why the data were being collected; they wished to know how many schools were being investigated and what the results were. They were friendly, cooperative, enthusiastic and were attentive despite the distractions in the environment described above. Their responses were articulate and they showed a better than average amount of verbal facility. Many of the P. S. children, on the other hand, seemed to be suspicious and hesitant. Their responses tended to be monosyllabic and in some instances it seemed that their attention was lagging.

Concluding Comment

The children in the school representing a relatively low socioeconomic level tended to hold the individual child responsible for any violation of rules of conduct and they thought in terms of punishment of the offending child to a much greater extent than did children representing a more favorable socioeconomic level. The latter children seemed to be able to comprehend a situational or environmental basis for misconduct. Therefore, they frequently suggested that the circumstances which produced the misbehavior be changed. As a characteristic response, the children of low socioeconomic status suggested the truant officer as an authority to punish the child who played hookey, whereas the children of higher status wondered if truancy might not be due to the fact that the child was unhappy at school and if so, a change of school might be the proper solution.

There were certain differences in the children's concepts of money. Children in both schools recognized the need for money but the children of low socioeconomic status suggested working after school to earn money by selling papers or shining shoes or the like while children of higher socioeconomic status spoke of allowances and none thought of working after school outside the home, although one child suggested making the beds in the morning as a means of adding to his funds.

The children of low socioeconomic status frequently referred to the exercise of legal authority. In commenting on the more serious offenses such as stealing and truancy, the children mentioned reform schools, homes for correction and truant officers. The children of higher socioeconomic status did not reveal familiarity with these guardians of law and order. It is interesting to note that one child in the latter group suggested that a psychiatrist be consulted.

The children in this study no doubt reflected the thinking and behavior patterns of the home. Perhaps the generosity of those in the higher socioeconomic group and their tendency to explain a child's misconduct may be due to their greater economic and emotional security. At any rate, the children who came from homes of a relatively low social, educational and financial level seemed more inclined to punish a culprit and to avenge his misdeeds than did the children who had been reared in homes of higher socioeconomic status.

THE EFFECT OF TRAINING ON THE TEMPO OF DEVELOPMENT OF GRAPHIC REPRESENTATION IN PRESCHOOL CHILDREN*

ELISABETH RUCH DUBIN
University of Chicago

THE PROBLEM

The central problem attacked in this investigation has been that of measuring the effects of a certain type of training on the speed of development of skill in graphic representation as studied through the medium of easel painting. An essential step in the development of artistic expression in the young child is learning that objects can be represented. This idea and some skill in carrying it out has been found to be acquired during the preschool years. It has been noted in an earlier study by Marian Monroe[1] that preschool children go through a typical sequence of development in learning to represent graphically. Our training methods were designed specifically to aid in speeding our experimental subject's progress through the following stages outlined by Monroe: (1) Scribble-unnamed, (2) Scribble-named, (3) Diagram, (4) Design, and (5) Representation. An attempt was made in the present experiment to minimize training in the more strictly aesthetic aspects of painting, the representative qualities being those on which attention was focused.

In order adequately to measure the effects of the training each experimental child was paired with an untrained control child. Thus, if measurably greater progress was made by the trained children in the time allotted to the experiment we would have reason to suspect that our training techniques had been successful in increasing our subject's rate of learning this particular skill.

Our problem is related to several other areas of inquiry, namely, those concerned with the question of (1) the development of learned and unlearned behavior, (2) the development of aesthetic and symbolic processes, and (3) pedagogical techniques in nursery education.

Both theoretically and methodologically this research is partially analogous to studies made by Munn,[2] Shirley,[3] Hilgard,[4] Dennis,[5] Gesell,[6] and Carmichael[7] who have each attempted to discover the degree of learnedness of various behavior items. Our experiment differs from most of the studies mentioned in that our control children were not trained at a later date than the experimental in order to determine the effect of greater maturity on speed of learning. Our finding, on the other hand, to the effect that training is influential in increasing the tempo of learning of a particular complex of skills and ideas which we term graphic representation is on the whole consistent with the findings of previous researches of a similar nature. What we were dealing with was what might be termed a behavior-pattern (using McGraw's[8] concept) which was found to be at a stage of relative modifiability through training at the preschool level. We measured the effect of a specific kind of training in speeding up the tempo of emergence of this pattern, and found the effect to be significant, at least for a few months immediately following the application of the training. Thus the present study finds its place in the more general fields of genetic and educational psychology.

In regard to teaching methods let us remark that we are in agreement with such per-

* An Article Based on the Essential Portions of a Ph.D. Dissertation in Human Development, the University of Chicago, 1946.

[1] Marian Monroe, "The Drawings and Color Preferences of Young Children", Unpublished Ph.D. Dissertation, the University of Chicago, 1929.

[2] N. L. Munn, *Psychological Development* (New York: Houghton–Mifflin Co., 1938), pp. 1–332.

[3] M. M. Shirley, "Locomotor and Visual-manual Functions in the First Two Years," *A Handbook of Child Psychology*, 2nd ed., ed. by Carl Murchison (Worcester: Clark University Press, 1933), Ch. V.

[4] J. R. Hilgard, "Learning and Maturation in Preschool Children," *Journal of Genetic Psychology*, XLI (1932), pp. 36–56.

[5] Wayne Dennis, "The Effect of Restricted Practice Upon the Reaching, Sitting, and Standing of Two Infants," *Journal of Genetic Psychology*, XLVII (1935), pp. 17–32.

[6] Arnold Gesell, *The Mental Growth of the Preschool Child* (New York: MacMillan Company, 1926), pp. x+447.

[7] Leonard Carmichael, "A Re-evaluation of the Concepts of Maturation and Learning," *Psychological Review*, XLIII (1935), pp. 450–470.

[8] M. B. McGraw, *Growth: A Study of Johnny and Jimmy* (New York: D. Appleton–Century Co., 1935), pp. xxi+319.

sons as Dewey,[9] Pearson,[10] Mearns,[11] and Cole[12] when they contend that the child cannot be left to solve all his artistic production problems alone, and that the teacher can successfully function as a catalytic agent in his development.

The Experiment

Our subjects were fifty-two preschool children divided equally between an experimental and a control group. The research was initiated with a larger number of subjects, but due to illness, uncooperativeness and some turn-over in the school population a number of pairs had to be dropped.

The method used in this study involved the setting up of parallel experimental and control groups with pairs from each of the classes at the University of Chicago Nursery School, i.e., from the two-year-old, younger three-year-old, older three-year-old, and four-year-old groups. The pair members in addition to being enrolled in the same nursery school class, were matched on the basis of sex, age and to a certain extent, drawing interest. There were two reasons for drawing the members of any given pair from the same nursery school class. One reason was practicality; but vastly more important was the second reason, namely, that this arrangement helped to prevent the experiential background of the children constituting any pair from being widely different, at least insofar as the nursery school situation was concerned. Sex was chosen as a basis for pairing because numerous researchers have found that sex differences occur fairly consistently in the art products of young children. We refer, for example, to studies by Goodenough,[13] Dennis,[14] Lehman and Witty,[15] Lark–Horovitz,[16] and McCarty.[17]

[9] John Dewey, et al., *Art and Education* (Rahway, N. J.: Barnes Foundation Press, 1929), pp. 180–183.

[10] Ralph Pearson, *The New Art Education* (New York: Harper Bros., 1941), pp. 207–214.

[11] Hughes Mearns, "Creative Spirit and Its Significance for Education," *Creative Expression* (Milwaukee: E. M. Hale and Co., 1939), pp. 13–21.

[12] N. R. Cole, *The Arts in the Classroom* (New York: John Day Co., 1940), p. 3.

[13] F. L. Goodenough, *The Measurement of Intelligence by Drawing* (Yonkers: World Book Co., 1926), p. 57.

[14] Wayne Dennis, "The Performance of Hopi Children on the Goodenough Draw-a-Man Test," *Journal of Comparative Psychology*, XXXIV (1942), pp. 341–348.

[15] H. C. Lehman and P. A. Witty, "Play Interests as Evidence of Sex Differences in Aesthetic Appreciation," *American Journal of Psychology*, XL (1928), p. 455.

[16] Betty Lark–Horovitz, "On Art Appreciation of Children: III Textile Pattern Preference and Study," *Journal of Educational Research*, XXXIII (1939), p. 34.

[17] Stella McCarty, *Children's Drawings* (Baltimore: Williams and Wilkins Co., 1924), p. 25.

There could be little doubt as to the desirability of pairing experimental and control children on the basis of age, since practically all research on the development of drawing in children points to rapid changes with increase in age. The reader is referred to the investigations by Rouma,[18] McCarty,[19] Goodenough,[20] as well as some by Gesell[21] who comment on the importance of the factor of neuro-muscular maturity as related to age in the evolution of drawing. The study by Monroe[22] from which the present research stems, resulted in the setting up of an age scale for measuring development of skill in drawing.

Since we were not able always to match pair members exactly in age, we permitted a maximum difference of four months. The mean absolute difference in age for our two groups of subjects was, however, only 1.07 months.

The interest of the children in painting was also taken into account in a rough way in pairing experimental and control subjects because motivation is an important determinant of learning rate, and ours was to be a study of learning. No precise measuring procedure was used for gauging interest, this trait being judged rather by observing the frequency with which each child asked to paint, the time he spent at the easel, and the enthusiasm and attention he displayed when art materials were presented.

The first step in the research consisted in the collection of paintings made spontaneously by all the children in the Nursery School during the free play period throughout the Autumn Quarter of 1942 (Pre-training Period). The art work in the Nursery School free-play period was carefully planned by the experimenter and the teachers to insure that a minimum of training was given at this time. Each child's comments concerning his paintings, along with his name and the date of execution of the painting were written immediately upon its completion on each picture drawn. At the end of the period the paintings were classified on the basis of a scheme which was a slight modification of Monroe's system, the major modification consisting in the omission of Monroe's "letters" category. The

[18] George Rouma, *Le Langage Graphique l'Enfant* (Paris: Misch et Thron, 1913), pp. 1 ff.

[19] McCarty, *op. cit.*, pp. 134–135.

[20] Goodenough, *op. cit.*, pp. 1–177.

[21] Gesell, *op. cit.*, pp. x+447.

[22] Monroe, *op. cit.*, pp. 1–16.

categories in terms of which the paintings were classified included the following: Scribble-unnamed, Scribble-named, Diagram, Design, and Representation. Let us describe these classes briefly. All paintings consisting of decorating the page with lines or blobs, or solid masses of color, if left unnamed, were classified in the category, scribble-unnamed. (2) The same types of paintings as the above, if given a name by the child, were placed in the category, scribble-named. (3) The drawings which were placed in the class, diagram, consisted of lines and masses of color having a definite relationship to each other which was explained by the child in the verbal report of the drawing. The depiction of spatial relationships rather than the details of appearance was the diagnostic feature of diagram drawings. (4) All drawings in which lines or masses were repeated in some rhythmical sequence or regular order were classified as design. (5) All drawings which looked like the object named by the child sufficiently to be recognized as that object by the experimenter were placed in the category, representation.

It was not difficult to classify a painting in accordance with the scheme just described. However, as a guard against any possible bias of the experimenter's, another competent person was called upon to check the experimenter's classifications. Disagreement between judges was rare, i.e., they occurred in approximately one percent of the instances.

At the end of the pre-training period the children in the Nursery School were paired, as we have described, according to nursery school class, sex, age, and interest.

During the Winter Quarter of 1943 (Training Period) the children in the experimental group were given training. A standardized procedure was organized. Each child in the experimental group was taken by the investigator at each training session to the experimental room where painting materials were arranged. The materials consisted of an easel, 16″ x 24″ news print paper, water soluble vegetable paints in four colors—red, blue, green, and yellow—and long-handled brushes one-half inch wide. This equipment is identical with that used at the Nursery School, and consequently was familiar to the children. Since the purpose of the training was to accelerate the child's development from one stage to another in the painting scale, the specific training method used in a given instance depended upon the stage of development that the child was in at the time of training. Once the child's level was recognized, training was directed toward the next higher stage. The child, after introductory pleasantries was asked to paint some pictures. Enthusiasm was evidenced by most subjects.

The following passages give a description of the framework of the formalized training procedure: The teacher waited until the child had completed his first picture, and then she said, "Tell me about your picture." As a result of his answer and the nature of the painting produced, the experimenter immediately classified the product. On the basis of this classification of the first picture produced the experimenter continued with one of the following series of comments. The sequence was repeated for each picture the child painted during the training session (usually not more than two or three pictures, a restriction necessitated because of time limitations), unless, of course, a variation occurred in the level of the child's work.

(1) When the first picture fell into the category, *Scribble*, the training was directed toward helping the child produce a picture at the scribble-named level, i.e., toward giving him the idea that the picture he painted might be about some object. Following is the series of comments the experimenter adhered to:

"Tell me about your picture."

If there was no response E continued "What does your picture look like?"

If there was no response E continued "Is it a kitty?"

If there was no response E continued "Does it look like a doggie?"

If there was no response E continued "Tell me about your picture."

Whether or not there was a response E concluded "If you look carefully at your picture you can see that it is about something."

(2) When the child's initial product was of the *Scribble-named* variety, the training was directed toward carrying him to the diagram level by the following method. This is the sequence of remarks used in this case:

"Tell me about your picture."

If the child had given a general name E continued "Tell me more about it."

If there was no response E continued "What is this part?" (pointing)

Whether or not there was an answer, E continued "What is this part?" (pointing to another place)

Whether or not there was an answer, E concluded "If you look carefully you can see that there are lots of parts to your picture, and they make something nice altogether."

The idea here was to give the child the notion of complex content as well as the relationship of parts to the whole and to each other. This was accomplished by the questions and comments put to the child by the experimenter.

(3) Should the child's product be a *Diagram*, the training was directed toward getting him to produce a painting that was a recognizable representation of something by the following method:

After the experimenter had asked the child to tell about the picture, and after the child had finished describing it the experimenter made a simple line drawing of whatever the child had been trying to paint. Then the experimenter said, "This is what a (e.g.) house looks like too, isn't it? Now what else are you going to paint?"

When the child had answered, the experimenter continued, "That is a fine idea. Be sure that it looks like a (e.g.) boat. I'm sure that you can make it look like a boat."

The purpose of the above was to improve the child's attempts at representation by giving him direct visual aids and a great deal of encouragement, but to avoid teaching him merely to copy.

(4) Since the product classified as *Design*, was in a sense a form of representation, the child was encouraged to achieve the more conventional recognizable forms merely by the following method:

After the child had completed a design painting, the experimenter said, "Now let's paint a picture about something. You tell me what you are going to paint a picture about."

If the child did not immediately respond with an idea E continued "It might be a house, or a lady, or a bear, or anything at all that you are interested in painting."

If the child preferred to continue painting designs, no further urging to represent was given.

(5) Since a product involving the recognizable representation of something showed that the child had captured the idea of graphic symbolism, the training at this level was directed toward encouraging the child not only to maintain, but also to improve upon the work which he had achieved. To this end the following method was applied:

After the child had finished painting and had described his picture, the experimenter would say

"That is one of the very nicest pictures you have ever made. I hope you will make one just as nice for me the next time."

Then the experimenter would call attention to interesting features of detail, proportion, color, composition, etc., depending upon the nature of the particular picture painted.

It will be noted that only at this point was the aesthetic quality of the painting involved in the instruction. The aesthetic quality did not, however, enter into the original classification of a painting as representation.

The formalized procedure which has just been described had naturally to vary somewhat for the sake of interest and in accordance with the conversation which the child made, but the basic pattern was maintained. A complete diary of each training session was kept so that exactly what was said by experimenter and child is a matter of record. These records indicate that the outline set forth above was adhered to closely, the incidental remarks occurring in addition to, rather than to the exclusion of, the basic pattern. The essential features of the guidance technique were unvarying.

During the Spring Quarter of 1943 (Post-training Period) training was discontinued and the paintings produced spontaneously in the art periods of the various groups were again collected. These were classified as were the pre-training paintings on the basis of the modified Monroe scheme.

After the paintings had been classified a "drawing" score was computed for each child. The system of scoring drawings devised by Monroe was borrowed. Let us turn to Monroe herself for a description of her method:

. . . some individuals within each group influence the total number of drawings to a greater extent than others. Some of the children who came to the easels frequently have a large number of drawings included in their group while others who came to the easels only a few times contribute but a small proportion to the total drawings. In order to avoid this weighting in the direction of a few individuals a method of scoring was adopted whereby each child could be assigned a numerical value in terms of his average developmental stage in drawing. In order to do this, the types of drawings were ranked in order of the age at which they appeared most frequently. The unnamed scribble drawings have their highest percentage at the youngest age group. This class was arbitrarily given the score of 1. The other types of drawings all show an increase with age. This increase is most rapid in the case of the scribble-named drawings which maintain a level from three years of age.This class as the second stage in drawing was given the score of 2. Diagrammatic drawings reach their peak at three and one-half years. This class was given the score of 3. Design, representation, and letters all increase simultaneously and have their greatest frequency at the oldest age group, four and one-half years. These classes were each assigned the score of 4. The score given to each class of drawings conforms as closely as can be determined to the order of its appearance in the development of drawing. Each child was then given a score in drawing by computing the average score of all his drawings.[23]

Statistical Analysis

Essentially there are five questions which must be answered by the experiment: (1) What is the relationship between the two groups, i.e., experimental and control, in their level of development in graphic expression at the start of the experiment? (2) What is the relation of the level of development for each group in the post-training period to its own level of development in the pre-training period? (3) Does the experimental group show a development between the two periods significantly greater than that of the control group? (4) What is the relationship between the two groups in their level of graphic expression in the post-training period? (5) Are there any factors revealed by the data, other than the training received by the experimental group, which may account for the observed differences?

It is desirable to determine whether the two groups matched on the basis of nursery school class, sex, age, and drawing interest had similar levels of development in graphic expression at the initiation of the experiment. Table I lists the children in each group, giving the background data on the basis of which they were matched, and each child's drawing score in the pre-training period. The control group has a mean drawing score of 1.52 while the experimental group has a mean drawing score of 1.70. The difference between these two means is 0.18 with a standard error[24] of $\pm$.07. The value of t is 2.57. This leads to the conclusion that the mean drawing scores were different at approximately the two-percent level of confidence,[25] prior to the training of the experimental group.

Now let us ask whether the improvement made in the post-training period by each group relative to its own initial mean drawing score was significant. If the training was successful, the experimental group at least should certainly show a significant improvement. If both should show a gain in skill one would expect the trained group to advance the greater amount.

Table II shows the scores in both periods together with the differences in the scores for each group. The control group improved its average score from 1.52 to 1.75, a difference of 0.23 points. The standard error of the distribution of this difference is $\pm$0.08, making the value of t 2.88. The improvement in the control group was, therefore, probably significant at the one-percent level.

The experimental group, on the other hand, improved from a mean score of 1.70 to a mean score of 2.48. The difference is 0.78 $\pm$ 0.10, the value of t being 7.8. The latter statistic indicates a highly significant difference. Thus the control group showed some improvement and the experimental group highly significant improvement, the trained children gaining almost three and one-half times as many drawing score points as the non-trained children.

[23] *Ibid.*, pp. 27–28.

[24] C. C. Peters and W. R. VanVoorhis, *Statistical Procedures and their Mathematical Bases* (New York: McGraw-Hill, 1940), p. 165.

[25] *Ibid.*, p. 173.

TABLE I

BACKGROUND CHARACTERISTICS AND DRAWING SCORES OF THE EXPERIMENTAL AND CONTROL GROUPS IN THE PRE-TRAINING PERIOD

Nursery School Class	Pair Number	Sex	Control Group		Experimental Group		Difference in Mean Drawing Scores of Pair Members
			Age*	Drawing Score	Age*	Drawing Score	
Two-Year-Olds	1	F	2- 8	1.1	2- 8	1.5	0.4
	2	M	2- 3	1.0	2- 4	1.1	0.1
	3	F	2- 0	1.0	2- 0	1.0	0.0
	4	M	2- 7	1.0	2- 6	1.0	0.0
	5	M	2- 3	1.0	2- 3	1.5	0.5
	6	F	2- 4	1.0	2- 6	1.6	0.6
Younger Three-Year-Olds	7	F	3- 3	1.6	3- 3	1.8	0.2
	8	M	3- 5	1.8	3- 5	2.0	0.2
	9	M	2-10	2.0	2- 8	1.0	—1.0
	10	F	2-10	1.9	3- 2	2.0	0.1
	11	M	2-11	1.9	2-11	2.0	0.1
	12	M	2-11	1.8	2-11	1.8	0.0
	13	M	3- 7	1.8	3- 6	2.0	0.2
Older Three Year-Olds	14	M	3- 8	1.6	3- 8	1.8	0.2
	15	F	4- 0	1.8	3-11	2.0	0.2
	16	M	4- 0	1.8	4- 0	1.4	—0.4
	17	M	3-11	1.4	3-11	1.4	0.0
	18	M	3- 9	1.5	3- 8	1.6	0.1
	19	F	3- 7	1.7	3- 9	1.5	—0.2
Four-Year-Olds	20	F	3-11	1.7	4- 0	2.3	0.6
	21	M	4- 6	1.6	4- 4	2.0	0.4
	22	M	4- 3	1.0	4- 0	1.4	0.4
	23	F	4- 7	1.9	4- 8	2.4	0.5
	24	F	4- 2	1.8	4- 6	1.9	0.1
	25	F	4- 2	1.7	4- 3	2.5	0.8
	26	M	3-10	1.2	3-11	1.7	0.5
Total				39.6		44.2	4.6
Mean			3- 3	1.52	3- 5	1.70	0.18
Sigma of Distribution				.35		.41	

*As of October 1, 1942.

The question as to the equality of the units with which we are dealing must be raised. Evidence bearing indirectly on this question will be given shortly, but first let us present another line of evidence bearing on the question of relative gains. The difference between the mean gains in the two groups is 0.55 (0.78 — 0.23), while the standard error[36] of the difference in mean gains is ±0.12. The fact that the value of t is 4.58, indicates that the inter-group difference is highly significant. The extent of the improvement made by the experimental group was significantly greater than that achieved by the control group.

There is one final step in the basic analysis of the data. The mean drawing scores in the post-training period are: Control group—1.75; experimental group—2.48. The difference between the mean post-training scores is 0.73 with a standard error[37] of ±0.059. The value of t is 12.4. The difference is obviously highly significant.

The question may well be raised at this point as to whether or not the initial advantage of a higher average score for the experimental group resulted only in apparent greater improvement, the units at the higher levels of the rating scale being perhaps of lesser magnitude than those at the lower. It is possible to check on this problem by determining whether any group of children who score high on a test of graphic expression show a greater improvement in score after a period of time than a comparable group of children who scored low initially. It is very simple to test this point from our data. Taking the ten children who at the start of the experiment in the pre-training period scored lowest in the control group, we find that on the average they improved their score 0.3 points in the post-training period. The ten control children who initially scored highest

[36] *Ibid.*, pp. 167–168.

[37] R. L. Thorndike, "Regression Fallacies in the Matched Groups Experiment," *Psychometrika*, VII (1942), p. 88.

TABLE II

DRAWING SCORES OF THE EXPERIMENTAL AND CONTROL GROUPS IN THE PRE- AND POST-TRAINING PERIODS

Pair Number	Drawing Scores						Difference between Experimental and Control Differences
	Control Group			Experimental Group			
	Pre-training	Post-training	Difference	Pre-training	Post-training	Difference	
1	1.1	1.5	0.4	1.5	2.4	0.9	0.5
2	1.0	1.6	0.6	1.1	2.5	1.4	0.8
3	1.0	1.2	0.2	1.0	1.8	0.8	0.6
4	1.0	1.4	0.4	1.0	2.2	1.2	0.8
5	1.0	1.3	0.3	1.5	2.8	1.3	1.0
6	1.0	1.0	0.0	1.6	2.2	0.6	0.6
7	1.6	2.0	0.4	1.8	2.4	0.6	0.2
8	1.8	1.6	—0.2	2.0	3.2	1.2	1.4
9	2.0	1.8	—0.2	1.0	2.3	1.3	1.5
10	1.9	2.0	0.1	2.0	2.6	0.6	0.5
11	1.9	1.9	0.0	2.0	2.3	0.3	0.3
12	1.8	1.7	—0.1	1.8	2.3	0.5	0.6
13	1.8	1.8	0.0	2.0	2.4	0.4	0.4
14	1.6	1.5	—0.1	1.8	1.5	—0.3	—0.2
15	1.8	1.7	—0.1	2.0	2.3	0.3	0.4
16	1.8	1.5	—0.3	1.4	2.3	0.9	1.2
17	1.4	1.3	—0.1	1.4	2.2	0.8	0.9
18	1.5	1.5	0.0	1.6	2.0	0.4	0.4
19	1.7	1.0	—0.7	1.5	3.2	1.7	2.4
20	1.7	2.4	0.7	2.3	2.1	—0.2	—0.9
21	1.6	2.7	1.1	2.0	3.7	1.7	0.6
22	1.0	1.6	0.6	1.4	2.6	1.2	0.6
23	1.9	3.4	1.5	2.4	3.1	0.7	—0.8
24	1.8	2.6	0.8	1.9	2.4	0.5	—0.3
25	1.7	1.7	0.0	2.5	2.9	0.4	0.4
26	1.2	1.9	0.7	1.7	2.7	1.0	0.3
Total...	39.6	45.6	6.0	44.2	64.4	20.2	14.2
Mean ..	1.52	1.75	0.23	1.70	2.48	0.78	0.55

improved their scores on the average only 0.15 points, just half as much in points as their low-scoring mates. The same relationship is true of the experimental group. The ten lowest scoring children improved on the average 1.15 points after receiving the training, while the ten highest members of the experimental group did only half as well, improving 0.59 points, on the average, after being trained. The evidence is strongly suggestive that the higher average initial score among the experimental children was not only not an advantage, but that the advantage lay with the group which measured low initially in the drawing score scale.

The problem of the effect on differences in group behavior of initial disparity in drawing scores we have attacked further by using a new pairing of experimental and control children. If we start with pairs of children who are matched exactly in initial drawing score, then we eliminate initial inequality as a factor in improvement rate. Table III shows twelve pairs of children who are matched exactly in initial drawing score. The mean difference between experimental and control subjects in the post-training period is 0.63 points in favor of the experimental group. The standard error of this difference is ±0.199. The ratio of the difference to the standard error of the difference is 3.18, a fact which indicates significance between the two-percent and one-percent levels of confidence and supports our other evidence to the effect that training of the type given tends to speed progress in ability to represent graphically.

It is of further interest to consider what happened to the variability of scores in the control and experimental groups between the two periods. The control group had a standard deviation for the pre-training period scores of ±0.35 while the standard deviation for the post-training period was ±0.53. This represents an increase in the score variability of 0.18. Since the standard error of the standard deviation[28] is ±0.084 the value of t in the case of the difference in variability in

[28] Peters and VanVoorhis, *op. cit.*, p. 182.

TABLE III

BACKGROUND CHARACTERISTICS AND PRE- AND POST-TRAINING DRAWING SCORES OF EXPERIMENTAL AND CONTROL GROUPS MATCHED ON THE BASIS OF PRE-TRAINING SCORE

Nursery School Class	Control Group					Experimental Group					Difference in Post-Training Scores
	Child Number	Sex	Age	Drawing Score		Child Number	Sex	Age	Drawing Score		
				Pre-Training	Post-Training				Pre-Training	Post-Training	
Two-Year-Olds	1	F	2- 3	1.1	1.5	2	M	2- 4	1.1	2.5	1.0
	3	F	2- 0	1.0	1.2	3	F	2- 0	1.0	1.8	0.6
	4	M	2- 7	1.0	1.4	4	M	2- 6	1.0	2.2	0.8
Younger Three-Year-Olds	8	M	3- 5	1.8	1.6	7	F	3- 3	1.8	2.4	0.8
	9	M	2-10	2.0	1.8	8	M	3- 5	2.10	3.2	1.4
	12	M	2-11	1.8	1.7	12	M	2-11	1.8	2.3	0.6
Older Three-Year-Olds	14	M	3- 8	1.6	1.5	18	M	3- 8	1.6	2.10	0.5
	16	M	4- 0	1.8	1.5	14	M	3- 8	1.8	1.5	0.0
	17	M	3-11	1.4	1.3	17	M	3-11	1.4	2.2	0.9
	18	M	3- 9	1.5	1.5	19	F	3- 9	1.5	3.2	1.7
Four-Year-Olds	20	F	3-11	1.7	2.4	26	M	3-11	1.7	2.7	0.3
	23	F	4- 7	1.9	3.4	24	F	4- 6	1.9	2.4	—1.0
Total				18.6	20.8				18.6	28.4	7.6
Mean Score				1.55	1.73				1.55	2.36	0.63
Sigma of Distribution											0.66

pre- and post-training periods is 2.14. This is probably a significant difference (at approximately the five-percent level of significance).

For the experimental group the standard deviation of the drawing scores was initially ±0.41, and in the post-training period ±0.46. The change in group variability is 0.05 ± 0.077, and hence is insignificant, the value of t being 0.65.

The increased variability noted in the control group is accounted for in part by a stretching of the distribution of the scores in the post-training period toward the high end of the scale. The increase in variability in the control group scores would, therefore, if anything, tend to favor this group in any comparison with the experimental group.

Another statistic of interest because of the possible light it may throw on the motivation in our two groups is the number of paintings the children produced. From the previous description of the method by which the drawing scores were obtained, it is evident that the score of each child was uninfluenced by the number of paintings he made. Consequently the number of paintings produced does not enter directly into the calculation of the group drawing score means. The control group increased the average number of its drawings from 9.8 in the pre-training period to 14.3 in the post-training period, or 4.5 drawings per child. The experimental group, on the other hand, better than doubled its number of drawings, showing an average gain of 13.3 drawings from an initial mean of 10.7. The difference between the mean gain in number of drawings of the two groups was 8.8 ± 4.5. The value of t in this case is 1.98. Thus, although the experimental group did show a decided increase in the number of pictures produced, this increase is significant only at the five-percent level.

CONCLUSIONS

It seems highly probable from the analysis of our data that the training of the sort given was effective in improving, by a significant degree, progress in the development of skill in graphic representation of our young subjects and also their interest in drawing. It also seems evident that our nursery children under a laissez-faire program over an average period of six months did not very conspicuously improve their drawing skill as a result of self-initiated learning.

This result indicates that easel painting involves a behavior-pattern sufficiently highly developed and yet still sufficiently malleable at the age of two years, as to be easily influenced by training. From this it would follow that in the nursery school situation it would be possible to develop an organized art program even for the two-year-olds which, while involving no negative effects on spontaneity and creativity in artistic behavior, could at the same time have a positive value in advancing children's drawing level.